How accurately can river height, width, slope and discharge be estimated from ideal SWOT outputs?

Senior Thesis

Submitted in partial fulfillment of the requirements for the

Bachelor of Science Degree

At The Ohio State University

By

Ethan O'Connor

The Ohio State University

2013

Approved by

Michael T. Durand, Advisor

School of Earth Sciences

Table of Contents

Abstract ii

Acknowledgements iii

List of Figures and Tables .... iv-v

Introduction 1

Study Area 2

Methods

Software Overview 3-4

SWOT Simulator Overview 5

Instrument Simulator Input Preparation 6-12

Simulator Output Preparation for Height and Slope Analysis 13-17

RivWidth Input Preparation for Width Analysis 18

Data Preparation for Discharge Analysis 18-19

Results

Height Error 20-29

Width Error 30-34

Slope Error 34-38

Discharge Error 39-42

Conclusion 43-44

Suggestions for Future Work 44-45

References Cited 45

Abstract

The proposed SWOT (Surface Water and Ocean Topography) mission is intended to produce, for the first time, global estimates of surface water storage change and discharge. With many advancements in remote sensing NASA is planning for this satellite to produce data products of high accuracy. In order to estimate the actual performance of the SWOT satellite a simulator was developed by NASA scientists at the Jet Propulsion Laboratory to produce SWOT outputs given specific inputs.

This study simulated SWOT observations over a study reach along the Sacramento River in California. The instrument simulator produces a cloud of point measurements of the river. Four methods of calculating reach-averaged river height, width, and slope from the point cloud data were explored. Errors in surface water height, width, and slope were characterized. These measurements are well within the science requirements. These values were used to estimate reach-averaged discharge along the study area, and calculated discharges were compared to modeled discharges to assess accuracy. Discharge mean absolute error values ranged from 6.6 % - 16.5 % depending on the method used to calculate SWOT measurements from the point cloud. Errors in calculating river discharge will be more heavily dependent on nontechnical factors rather than instrument capability.

Acknowledgments

The completion of this work would not have been made possible without support from the following. This support came in all forms to include; scientific advice, technical aid, mentoring, occasional jokes, and I cannot forget to mention the free pizza on Fridays. All of this support was well received and much appreciated.

To my thesis advisor, Professor Michael Durand, thank you for the opportunity to be involved with your research. Despite an incredibly demanding schedule, filled with other projects and graduate students, you always made time to review my work and point me in the right direction. Thank you for your constant effort and encouraging ideas.

To my mentor and colleague, Rodrigo Cauduro Dias De Paiva, thank you for sharing your experience and your workspace with me. I find it hard to believe you get any of your own work done with my constant nagging, but it is safe to assume that without your help I would be stuck reading through the ArcGIS help manual instead of finishing this project. Thank you.

To my friend and coworker, Reza Najafi, thank you for your time and encouragement during your time at The Ohio State University. Hearing your stories about when you first started doing research always put me in the right frame of mind when I was stuck on my project. You contributed heavily to the friendly atmosphere of our workspace, and made it enjoyable to work there.

List of Figures and Tables

Figure 1: A large scale view of the study area, including California, the coast line and the Ocean

Figure 2: A small scale view of the study area, showing the clear upper and lower limits of the study reach and the river morphology

Figure 3: Original HEC-RAS cross sections, used to create three-dimensional river bathymetry

Figure 4: TIN created from original HEC-RAS cross sections

Figure 5: Revised HEC-RAS geometry with interpolated cross sections

Figure 6: Tin created from revised HEC-RAS cross sections

Table 1: A table used to compare data products to their source, HEC-RAS cross sections, to ensure each product agrees with the others

Table 2: This table shows the difference between the modeled HEC-RAS water surface and the interpolated HEC-GeoRAS surface. It also shows the difference between the terrain bathymetry and simulated flood depth, compared to HEC-GeoRAS surface

Figure 7: Flow chart depicting the steps required to prepare SWOT simulator inputs

Figure 8: Point Cloud

Figure 9: Point Raster

Figure 10: Bounding Polygon Raster

Figure 11: Channel Limited Raster

Figure 12: Buffered Point Raster

Figure 13: River Raster

Table 3: Height data statistics for each of the methods used to analyze SWOT outputs

Figure 14: Point raster height error histogram

Figure 15: Bounding polygon raster height error histogram

Figure 16: Channel limited raster height error histogram

Figure 17: Buffered point height raster error histogram

Figure 18: River raster height error histogram

Figure 19: Point raster height error map

Figure 20: Bounding polygon height error map

Figure 21: Channel limited height error map

Figure 22: Buffered point height error map

Figure 23: River raster height error map

Figure 24: Comparison of the bounding polygon raster width outputs to reference data

Figure 25: Comparison of the channel limited raster width outputs to reference data

Figure 26: Comparison of the buffered point raster width outputs to reference data

Figure 27: Comparison of HEC-GeoRAS to HEC-RAS modeled data.

Figure 28: Depth profile for the study area

Table 4: This table lists slope results and their corresponding source.

Figure 29: Slope profile generated from the point raster

Figure 30: Slope profile generated from the bounding polygon raster

Figure 31: Slope profile generated from the channel limited raster

Figure 32: Slope profile generated from the buffered point raster

Table 5: Calculated reach average discharge for HEC-RAS data

Table 6: Table displaying all reach-averaged discharges data for different raster techniques

Table 7: Errors in discharge of different filtered data sets

Table 8: Errors in discharge, of filtered data sets calculated with true HEC-RAS slopes

Table 9: Errors in discharge, of filtered data sets calculated with true HEC-RAS widths

Introduction

Surface fresh water is a vital resource with many uses, none of which are more important than sustaining life. However, it is alarming how little of an understanding there is concerning the spatial and temporal dynamics of surface water storage and discharge. River discharge as well as lake and wetland storage of water are critical terms in the surface water balance, yet they are poorly observed globally and the prospects for improvement from in-situ networks are bleak [*Shiklomanov et al,. 2002*]. Even current climate models fail to provide an accurate estimate of water storage change, producing errors between 50 and 100% [*Roads et al,. 2003*]. This severe lack of understanding and prediction of surface water storage change is not only disappointing, but also dangerous especially with the constant concern of climate change.

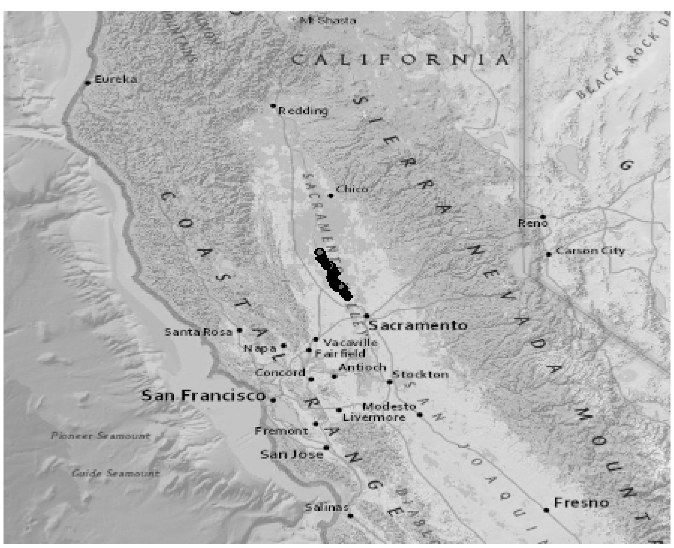
To address this problem, a spaceborne mission has been undertaken by NASA. The Surface Water and Ocean Topography (SWOT) mission will include an accurate measuring radar based system that can reliably estimate surface water storage changes by reporting surface water elevation for points it classifies as water inside of an observation swath. These observations can then be used to estimate discharge. The project is required to assess water bodies with area greater than $(250\text{m})^2$ and rivers of average width greater than 100 m but also includes a science goal of measuring rivers of average widths of 50 m.

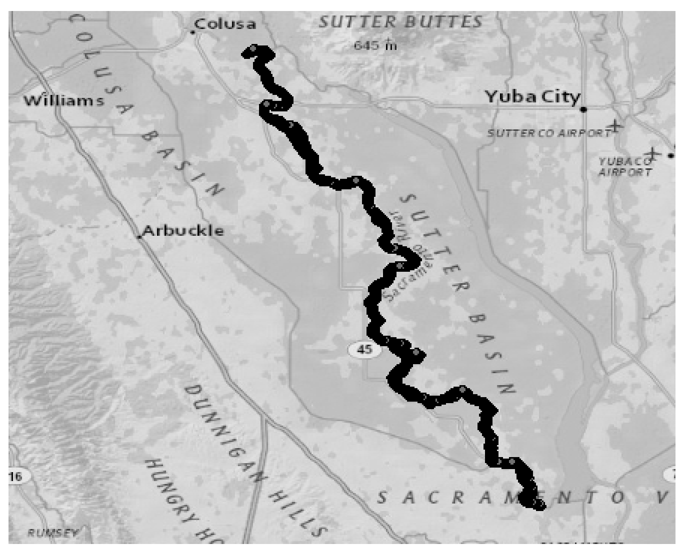
This study will produce simulated SWOT observations by utilizing a SWOT Simulator designed by NASA's Jet Propulsion Laboratory. This simulator produces latitude, longitude, and elevation outputs for user supplied areas of interest based on the proposed SWOT orbit.

The primary goal of this study is to characterize the error in height, width, and slope measurements obtained from SWOT outputs by comparing them to modeled data, using the SWOT instrument simulator. The secondary goal of this study is to estimate discharge along the study area, by using these height, width, and slope measurements. Currently there is no published literature using SWOT Simulator outputs in order to estimate river discharge.

Study Area

This project focuses on the Sacramento River, specifically the region between the cities of Colusa and Sacramento. The upper limit of the study reach corresponds to the physical extents of an existing hydraulic model that was obtained from the USGS. The lower limit was taken where the river widened considerably to the low lying floodplain near the Yolo Bypass. The study reach is 82,475.6 meters long, and is shown in Figures 1 and 2 below.





Figures 1 and 2**:** The portion of the Sacramento River studied, represented in black by the simulator outputs.

Methods

Software Overview

Three main tools were used in order to prepare the appropriate inputs required by the SWOT simulator. HEC-RAS was used to perform one-dimensional simulations of the behavior of the water flow through the Sacramento River. ArcGIS and HEC-GeoRAS were used to create three-dimensional maps of the water surface elevation. Each of these tools are briefly discussed and described.

Hydraulic Flow Modeling

The Hydrologic Engineering Center's River Analysis System (HEC-RAS) performs one-dimensional river analysis for steady and unsteady flow conditions. This software contains a graphical user interface used to model river systems by creating cross sections of the river channel with referenced elevation data. Also HEC-RAS contains data management capabilities that allow the user to test multiple scenarios on the same river geometry, and reporting functions to view and export simulation results.

An existing HEC-RAS model of the Sacramento River Basin produced by the Army Corps of Engineers and the United States Geological Survey was used to provide detailed information on river bathymetry, since each input cross section is surveyed to ensure accurate modeling. HEC-RAS was also used to produce cross section specific flow outputs to include; water surface height, top width, slope, and discharge, calculated from user defined flow conditions. For this study, one 22-day flood event was simulated in HEC-RAS.

Mapping and Spatial Analysis

ArcGIS developed by ESRI, is a geographic information system, designed to allow people the ability to organize, manage, merge, and analyze geographic information. This software allows the user to manipulate input data sets and create new data by utilizing different tools in order to address spatial analysis questions.

ArcGIS was used in this project to combine HEC-RAS data with a digital elevation model (DEM) of the study area. It was also used to visualize and analyze SWOT simulator outputs, and to produce the input data required by RivWidth.

3-Dimensional Flood Plain Modeling

The Hydrologic Engineering Center's geospatial extension to the River Analysis System (HEC-GeoRAS), designed to allow the integration of ArcGIS and HEC-RAS data. This extension is used inside of ArcGIS and allows users with limited GIS experience to create and export geometry to HEC-RAS for flow analysis. It can also be used to import HEC-RAS outputs and convert them to geospatial data.

HEC-GeoRAS was used to export an existing HEC-RAS model and flow simulation results of the study area to ArcGIS and then used for floodplain mapping. Floodplain mapping projects a three dimensional water surface over an underlying DEM to show water surface extent on the terrain.

Measuring River Width

RivWidth is a program that is designed to calculate the width of a river. It does this by first producing a modeled centerline, and then determines the distance from that centerline to the water's edge. The only input required by this program is a classified raster that contrasts inundated pixels and land pixels.

SWOT Simulator Overview

The SWOT simulator is a Linux-based tool that mimics the expected performance of the SWOT satellite currently in development. The main instrument onboard SWOT is a radar interferometer. The interferometer emits radar pulses focused on the same ground target area from two radar antennae. The pulses are reflected off of the Earth and are recorded by the interferometer. The phase difference between the two radar return pulses allows the interferometer to construct an interferogram, the interference pattern between the two radar pulses. User inputs for the simulator include a SWOT orbit pass, a digital elevation map (DEM) of a scene of interest, and a depth DEM for water found in that scene.

The simulator consists of several modules that are run sequentially, often using one module's product as the input for the next module. Generally the simulator can be broken down into three main parts; geophysical data processing, interferogram formation, and height reconstruction. The first part processes the user supplied inputs and interpolates them to the cross-track and along-track geometry of the swath. In essence, selecting only the parts of the user input DEMs that the satellite will be able to see during the input orbit pass, and combining them to one image. It also produces a sensor file containing location and orientation information for the satellite required for height reconstruction. The interferogram formation module uses the outputs from the previous section to calculate interferograms; phase differences between the returned radar signals from a simulated SWOT overpass. Further processing allows the addition of along track point target response times. Height reconstruction is completed by using the interferogram to produce data points with associated heights. Adding thermal noise, error generated by the instrument, is made possible by combining a reference, no noise interferogram, with a noisy interferogram. The modules that comprise the Height reconstruction tool are designed to filter points affected by layover.

Each of these processing steps utilizes several independent modules, each requiring a parameter file to run. All input and output files used in these processes are in netCDF format, minus the user supplied DEMs. These DEMs can be any raster file that the Geospatial Data Abstraction Library (GDAL) recognizes.

Instrument Simulator Input Preparation

The SWOT simulator requires three user specified inputs. The first input is a DEM of the scene of interest, to include river bathymetry. The second input is a water depth DEM, containing depth information corresponding to each pixel in the scene DEM. The third input is an orbit file containing latitude, longitude, altitude, time, and heading data for the satellite.

To create the first input a 90 meter resolution DEM of California was obtained and manipulated in ArcGIS. This DEM was from the Shuttle Radar Topography Mission data set, and although captured land elevation, this DEM had no information about river bathymetry. Instead it captured the water surface elevation at the time of the satellite observation. To address this, the surveyed river bathymetry from the existing HEC-RAS model of the Sacramento River Basin was exported into ArcGIS and combined with the California DEM, resulting in a DEM with dry rivers. The DEM was resampled to 10 meter resolution and converted to the coordinate system of the HEC-RAS model; this ensured that latitudinal, longitudinal, and vertical datum agreed between all data sources. A shape file of the HEC-RAS model was imported to the map in

ArcGIS to clip unnecessary data away from the DEM; this not only saved hard drive space but also allowed ArcGIS to run much smoother and faster.

A HEC-RAS flow simulation was run to produce hydraulic data outputs for a 100 year flood simulation designed by USACE for flood modeling. These outputs were then exported from HEC-RAS, converted to the appropriate file type and imported into ArcGIS using the HEC-GeoRAS extension. The HEC-GeoRAS import tool adds multiple features to the ArcGIS map, containing various HEC-RAS data, one of which is called XS Cut Lines 3D (Figure 3). This feature contains elevation data for each cross section in the HEC-RAS model. An interpolated triangulated irregular network (TIN) surface that simulated the bathymetry of the HEC-RAS model (Figure 4) was produced using the cross sectional data. A smooth river channel was expected, but instead the TIN was very rough and distorted especially when the river turned quickly or was very thin. This was because there were too few cross sections in the original HEC-RAS model especially around the bending parts of the river. The original HEC-RAS geometry was then interpolated between the surveyed cross sections (Figure 5). The updated geometry produced a usable output of the dry-earth or bathymetry TIN (Figure 6).

The bathymetry TIN was then converted to a raster data set, which allowed me to replace the original DEM's inundated area with dry bathymetry data. The last step involved converting the coordinate system and vertical datum reference to agree with the SWOT simulator. This completes the process to prepare the first SWOT input.

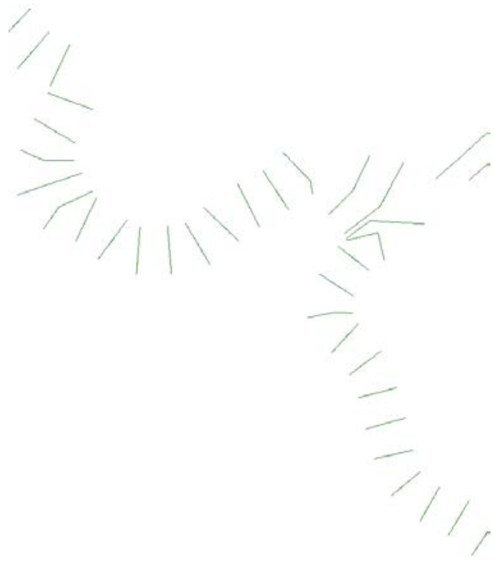


Figure 3: Original HEC-RAS cross sections, used to create three-dimensional river bathymetry.

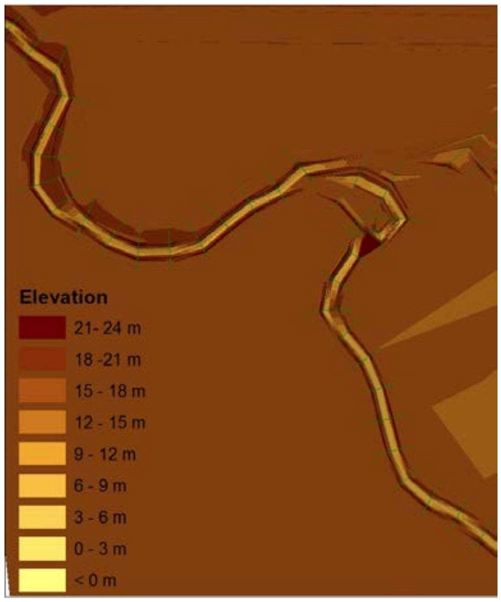


Figure 4: TIN created from original HEC-RAS cross sections.

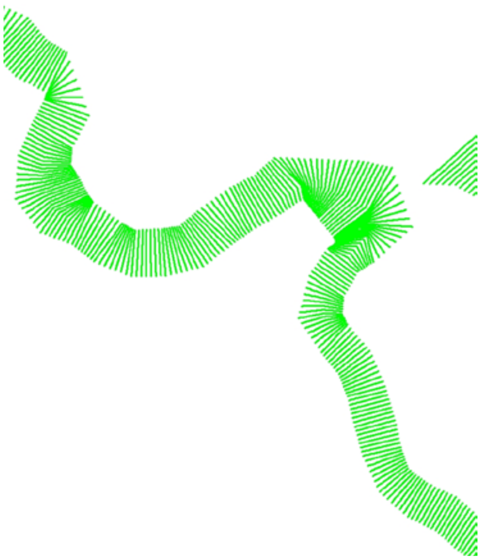


Figure 5: Revised HEC-RAS geometry with interpolated cross sections.

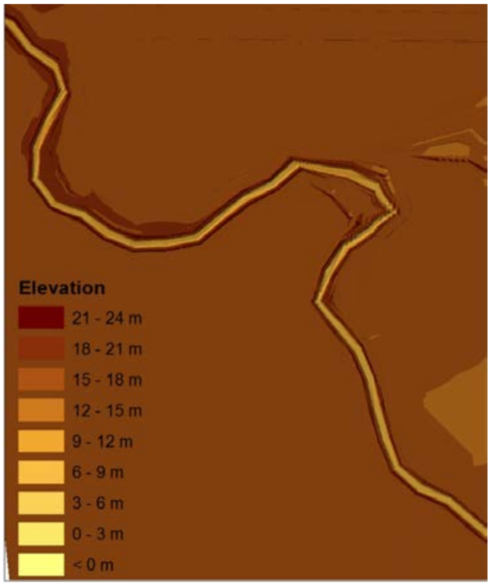


Figure 6: Tin created from revised HEC-RAS cross sections.

Producing the second input, the depth DEM involves far fewer steps, because this is the main output produced by HEC-GeoRAS. The original (pre-interpolation) HEC-RAS geometry file was used to proceed with flood plain mapping in order to avoid introducing any unexpected errors into the model which could influence its outputs, and translate those errors during floodplain mapping.

In order to import data into ArcGIS for floodplain mapping, HEC-GeoRAS requires the user to establish a terrain layer to project the water onto. Naturally the revised DEM that included the river bathymetry was used. The first step HEC-GeoRAS takes after importing HEC-RAS data is to generate a water surface elevation TIN. This TIN is a surface that forms between each cross section from HEC-RAS, based off of the modeled water surface height from the flow simulation. This water surface TIN is then compared to the terrain DEM. If the terrain DEM elevation is higher than the water surface TIN elevation, that pixel remains dry; otherwise the pixel becomes inundated with a depth calculated from the difference between the water surface TIN and the underlying terrain DEM. The water surface TIN, the depth raster, and bathymetry raster data was compared to HEC-RAS cross section data, to ensure there were no contradicting outputs (Table 1). The evaluated cross sections were randomly chosen throughout the study area, identified by the HEC-RAS cross section number. Ideally, the cross section water elevation should equal the water surface TIN elevation. Also the sum of the DEM bathymetry elevation and flood plain depth should also equal the cross section water elevation. Unfortunately this is not the case, and the errors are listed below (Table 2). Although the interpolated HEC-GeoRAS surface does not match exactly with the HEC-RAS surface, on average it only varies by 1.5 meters. However the correlation between the depth DEM and bathymetry DEM is much stronger, with an average difference of .12 meters. The depth raster was then geographically aligned to the SWOT simulator datum referencing scheme. This completes the process to prepare the second SWOT simulator input.

Sacramento Colusa-Feather Cross Section	Cross Section Water Elevation (m)	Water Surface TIN Elevation (m)	DEM Bathymetry Elevation (m)	Flood Plain Depth (m)
139	20.14	18.82	7.62	11.20
128	17.99	16.81	8.84	8.58
118.5	16.18	14.87	13.41	1.46
108	14.75	13.23	3.96	9.27
88	13.47	11.28	9.14	2.13

Table 1: A table used to compare data products to their source, HEC-RAS cross sections, to ensure each product agrees with the others.

HEC-RAS Cross Section	Difference between HEC-RAS and HEC-GeoRAS elevation (m)	Difference between Bathymetry + Depth and HEC-GeoRAS (m)
139	1.32	0
128	1.18	.61
118.5	1.31	0
108	1.52	0
88	2.19	-.01
Average	1.50	.12

Table 2: Difference between the modeled HEC-RAS water surface and the interpolated HEC-GeoRAS surface. Also shows the difference between the terrain bathymetry and simulated flood depth, compared to HEC-GeoRAS surface.

The third SWOT simulator input is the orbit file. Orbit files were provided for a test scenario on the Ohio River with the simulator software. These seven orbit files represent passes the satellite will take over Ohio when launched. In order to create orbit files to run the simulator over the Sacramento River Basin, the orbit longitude was shifted to be over California. Only one orbit pass used to created SWOT simulator outputs for this study. This completes the preparation for the third SWOT input.

A flow chart outlining the preparation of inputs for the SWOT simulator follows on the next page.

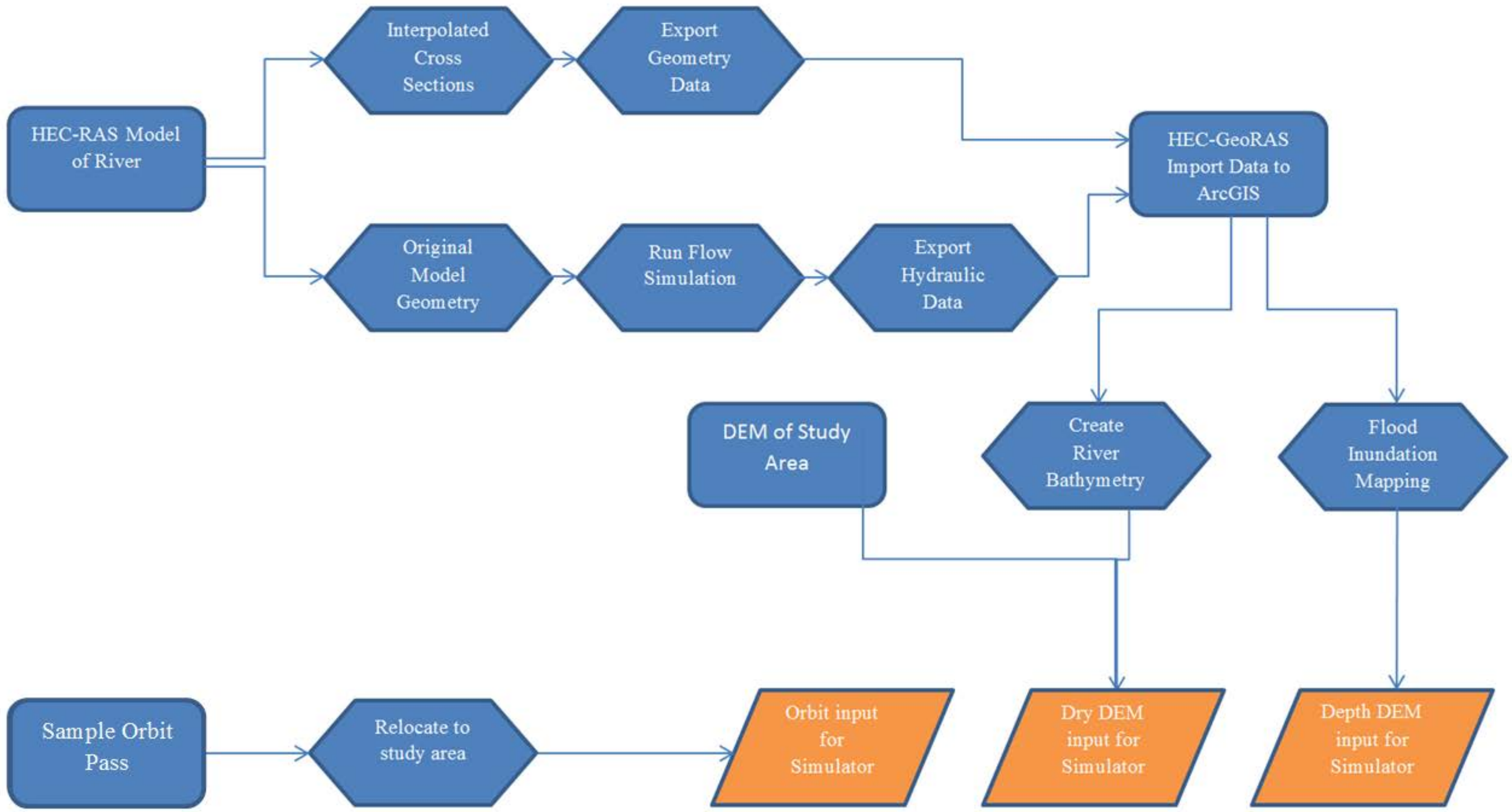


Figure 7: Flow chart depicting the steps required to prepare SWOT simulator inputs

Simulator Output Preparation for Height and Slope Analysis

The main output from the SWOT Height Reconstruction programs is a netCDF file containing latitude, longitude, and elevation data for SWOT measurements. This data was imported into ArcGIS and displayed as a point cloud (Figure 8). Point cloud data is difficult to analyze in ArcGIS so five different experimental conversion methods were used in order to produce raster data sets, pictured below in grayscale to show the differences in geometry each technique produces; the height errors associated with each geometry are discussed later.

The first method created a so-called "point raster" (Figure 9). This raster was created with a resolution of 10 meters and received its elevation data from the average of any simulator output that fell within a spatially corresponding pixel boundary. The majority of point cloud outputs are at least 10 meters apart so most raster pixels represent a single simulator output. Pixels that did not have any simulator points within their boundaries remained as empty raster cells containing no data. This method was intended to provide a raster data set that was closest to the output values of the simulator. There was no interpolation used in this method and therefore all data came directly from the simulator.

The next four methods involve different techniques to create raster data sets that interpolate the sparse point cloud to a continuous surface. First an interpolated surface was created using the simulator outputs. This raster was created with a resolution of 10 meters, and produced values for every pixel in the rectangular extent of the simulator output data frame. The challenge then became creating geometry that resembled the river to clip the interpolated elevation surface down to in order to evaluate the performance of the SWOT simulator. The second method created what can be referred to as a "bounding polygon raster" (Figure 10). A line was digitized between neighboring points that lay farthest from the center of the point cloud mass. Once this line wrapped completely around the simulator output points it was then converted to a polygon. This polygon was used as a boundary for the interpolated elevation surface. Only values that resided inside of this polygon were kept, all others were discarded.

The third method used a HEC-GeoRAS output, "Bounding Polygon" as the boundary for the interpolated elevation surface and created a "channel limited raster" (Figure 11). This polygon represents the channel limits from the HEC-RAS model. SWOT is intended to be a global mission and the vast majority of the rivers observed will not have a reference HEC-RAS model to obtain this Bounding Polygon from. However this technique was based on the idea of having satellite imagery of the observed river readily available. A smooth polygon can readily be traced over the channel of a river image and used to constrain the interpolated elevation surface produced by a point cloud of SWOT outputs. All elevation data outside of this polygon was discarded.

The fourth method created a "buffered point raster" (Figure 12). This raster was created by adding a circle, 100 feet in diameter, around each point in the point cloud. Overlapping circles were then merged together to create a continuous geometry. This buffered geometry was used as the boundary for the interpolated elevation surface from the point cloud. All data that fell outside of this geometry were discarded.

The fifth method created a "river raster" (Figure 13). This method was used to investigate the effects of applying a water mask to the simulator output points. In this case, a polygon was made from the floodplain depth file produced from HEC-GeoRAS. This polygon has the exact geometry of the depth DEM that was provided as an input to the SWOT simulator and represents the exact water surface extent throughout the study area. The polygon was then used as the boundary for the interpolated elevation surface, and all data outside of this geometry was discarded.

To determine the error in height from the SWOT simulation, a truth reference raster data set was needed. This was created by combining the depth DEM and the California with bathymetry DEM that were given to the SWOT simulator as inputs. The combination of these two DEMs produced a DEM that maintained the land elevation data, and replaced the bathymetry data with water surface elevation produced by HEC-GeoRAS. This DEM represents the elevation data the simulator registered. Each raster data set produced from the methods described above was then compared to this reference DEM to determine the error in SWOT height measurements.

In order to translate these height data into slope data, a few more steps are required. First the centerline of the river was digitized in ArcGIS to determine the length of the study area. Then the centerline was converted into a point feature. A channel polygon was digitized around the main channel of the river, based on the depth DEM. This polygon and the centerline points were then used to create zones based on Euclidean distance from the centerline point to the boundary of the channel polygon. These zones were used to extract height values from each of the height raster data sets previously produced, and assigned the average of each zone to its corresponding centerline point. These average centerline point heights were then plotted versus distance and a linear trend line was produced to obtain slope data.

To determine the error in slope from the SWOT simulation, the combined depth and California with bathymetry DEM was used as a truth reference and its slope data was compared to the slope data created from each height raster.

Figure 8: Point Cloud

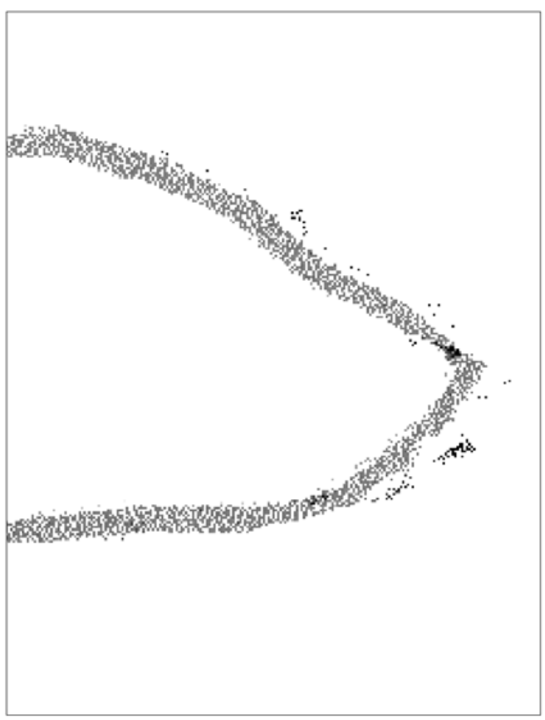


Figure 9: Point Raster

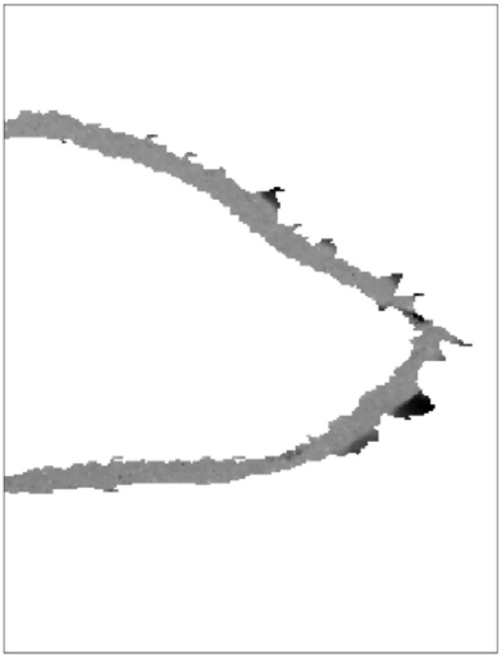


Figure 10: Bounding Polygon Raster

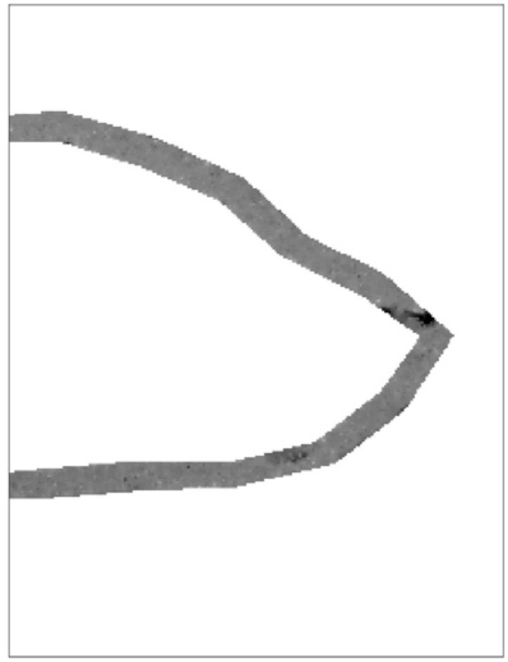


Figure 11: Channel Limited Raster

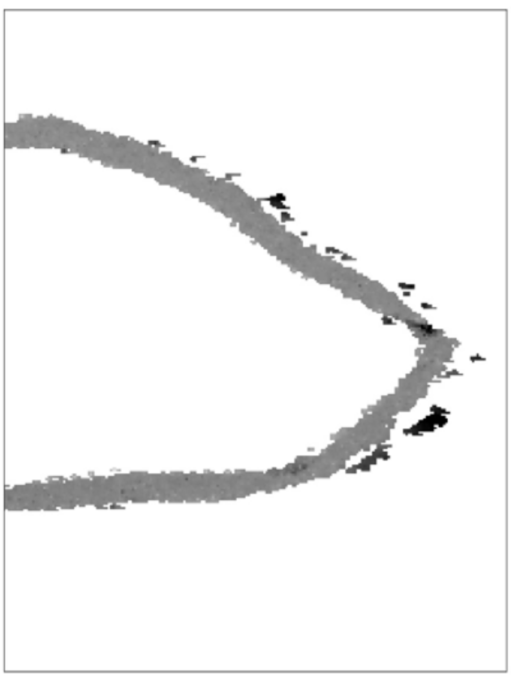


Figure 12: Buffered Point Raster

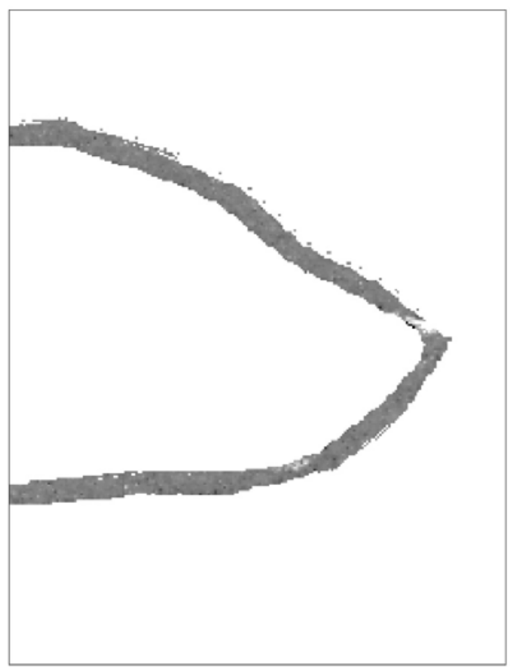


Figure 13: River Raster

RivWidth Input Preparation for Width Analysis

RivWidth requires one solitary input, a raster data set with two classifications; pixels are identified as either water pixels or land pixels. To create these inputs, the pixels in river geometry produced from each of the five methods above were reclassified. For each method, the area inside the river geometry was classified as water pixels and the area outside of the geometry was classified as land pixels. Originally these raster data sets were created with 10 meter resolution, but when they were input into RivWidth, errors refused to allow the program to finish its calculations. Once the data sets were resampled to 20 meter resolution, RivWidth ran without error, and produced width results.

To determine the width error from the SWOT simulation, the RivWidth output for the reclassified river raster was used as truth reference. This output was used as a reference because this data set only contained inundated area from the floodplain output produced by HEC-GeoRAS. The other four raster data sets used to calculate height and slope data were input into RivWidth and compared to the river raster width data. The HEC-GeoRAS floodplain was also compared to the cross sections widths from the HEC-RAS model to ensure they agreed.

Data Preparation for Discharge Analysis

After producing results for an estimated river slope and width, discharge could then be calculated for various sub reaches along the study area. The discharge estimations are calculated using Manning's Equation (Equation 1) shown below.

$$Q = \frac{1}{n} A^{\frac{5}{3}} W^{-\frac{2}{3}} S^{\frac{1}{2}}$$

Where Q is discharge (m^3/s), n is a manning's roughness coefficient (unitless), A is cross-sectional area (m^2), W is width (m) and S is slope (unitless).

The study reach was then subdivided into five reaches with length close to 10,000 meters, starting at 550 meters downstream of the study area. The whole study area was not used because the produced width results encounter errors and discrepancies around 57,000 meters flow distance from the upstream boundary, and these errors would propagate errors into the resulting discharges and produce meaningless data. The sub reaches were planned to be 10,000 meters in length, but in order to provide more accurate results, reaches were based on cross sectional data from in HEC-RAS. The first reach spans from 550 meters to 10,200 meters. The second reach spans from 10,600 meters to 20270 meters, the third reach spans from 20600 meters to 30300 meters, the fourth reach spans from 30700 meters to 40350 meters, the fifth reach spans from 47000 meters to 50800 meters. The reaches were based off of the HEC-RAS model because reach averaged cross sectional area were used in the discharge estimations, as well as the manning's roughness coefficient.

For each of the five sub reaches the slope data was recalculated. This allowed for more local accuracy for each sub reach. Then reach averaged widths were computed. Discharge was then calculated for each of the five sub reaches and compared to the reach average discharge of the HEC-RAS model. This process was repeated for each of the data rasters that were able to produce width outputs; the bounding polygon data set, the channel limited data set, the buffered point data set and the river raster data set.

Results

Height Error

The average errors and standard deviation for each of the five height data sets created are organized in Table 3 below. The histograms are Figures 14 through 18 below, and relate to the point raster, bounding polygon raster, channel limited raster, buffered point raster, and river raster respectively. The method that produced the best results, only a 2 centimeter error average, was the river raster. The second best results came from the point raster which produced height errors with an average of 10 centimeters. The worst results, 22 centimeter error average, came from the bounding polygon raster. The buffered point raster, and channel limited raster are relatively equal in quality, with errors ranging between 16 and 17 centimeters. The mean height error is non-zero (bias), and the error standard deviation of one meter is greater than expected due predominantly to geolocation errors of the data.

The majority of height errors with large magnitudes for all techniques occur in one of three regions. These regions include sections where the river becomes very thin, along the banks of the river but more so the right bank, and any point outside of the river. This occurrence agrees with the statistical data in Table 3. The height raster with the most pixels outside of the river channel is the bounding polygon raster, followed by the buffered point raster, limited channel raster, and point raster. The river raster only includes the input water surface area, which means it does not incorporates pixels from the SWOT simulator that are outside of the true river geometry. This drastically increases this data set accuracy. Why the right bank produces more error than the left bank can possibly be explained by the simulated physical positioning of the satellite. For the input orbit provided, the study area resides to the left of the satellite. This means that the right bank is closer to satellite and would be more affected by layover, height error in remote sensing products caused by the topography of the earth, than the left bank. The study area does not contain drastic changes in topography, such as mountains, however the river does have right and left levees that are steep and thin. These levees can cause simulator outputs to fall outside of the river geometry, due to the observation angle of the satellite, and to appear as though they are higher in elevation, due to difference in distance of the radar measurement, than the river.

The height error produced in the thin sections of the river is dependent on the radar capabilities of the SWOT payload. The instrument is rated to about 3 centimeter accuracy for 50 meter pixels. Areas where the river is barely 50 meters in width are unable to be observed correctly by the SWOT simulator.

These individual measurement errors for the point raster (Figure 19) are made worse by the other techniques (Figure 20, 21, and 22). This is because they are including area outside of the river channel that is receiving elevation data from the interpolated water surface, the majority of which are water surface values, and comparing those values to the known land elevation of the supplied DEM of California. In other words these techniques take elevation data from a weighted average of SWOT river elevation measurements, and compare them to land elevation data.

The most important results from the height error comparisons are the river raster height errors (Figure 23). This data set shows that the simulator produces very accurate results when only incorporating SWOT measurements known to occur on the water surface. The use of a water mask or other filtering technique can improve accuracy of height measurements by 80 to 90% for the mean error and by more than 50% for the error standard deviation.

Method	Average Error (m)	Standard Deviation (m)
Point Raster	.103	.99
Bounding Polygon Raster	.217	.97
Channel Limited Raster	.158	.85
Buffered Point Raster	.169	1.13
River Raster	.020	.45

Table 3: A comparison of data statistics for each of the methods used to analyze SWOT outputs

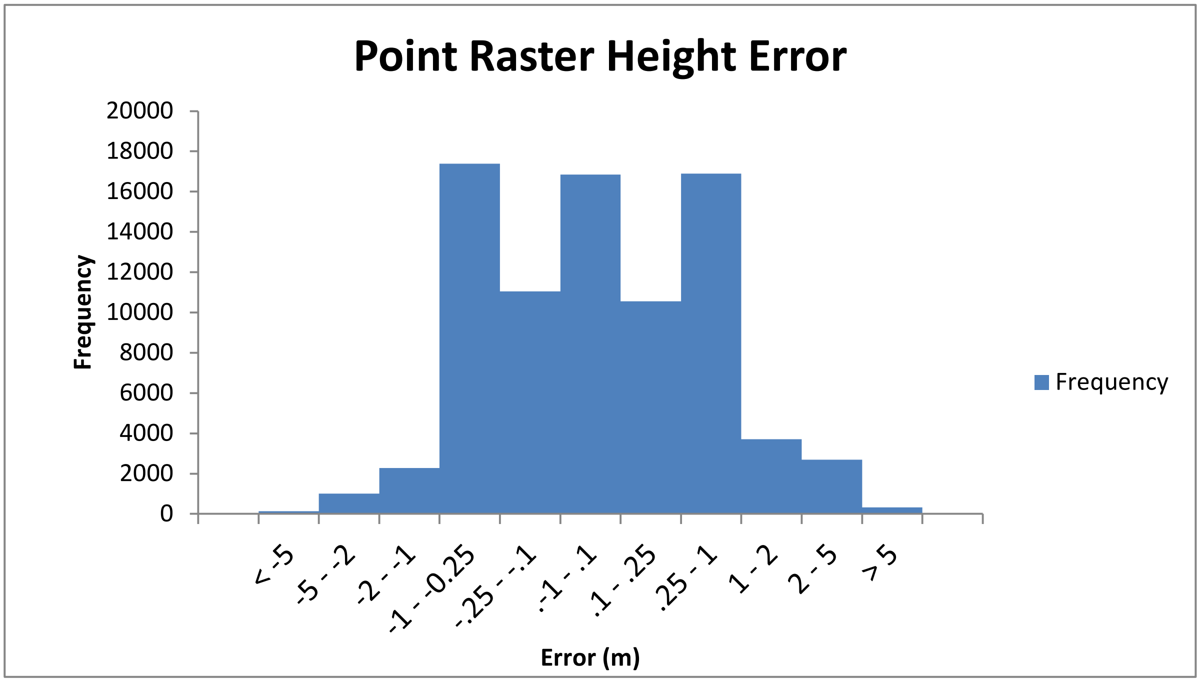


Figure 14: The point raster height error histogram

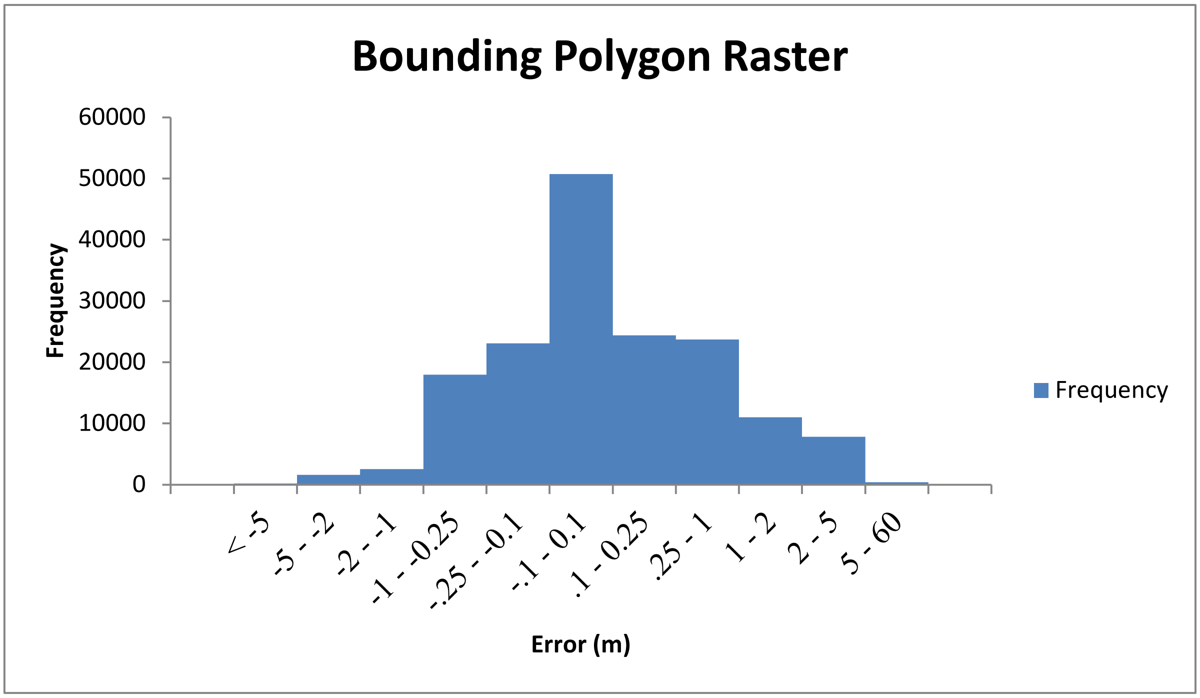


Figure 15: The bounding polygon raster height error histogram

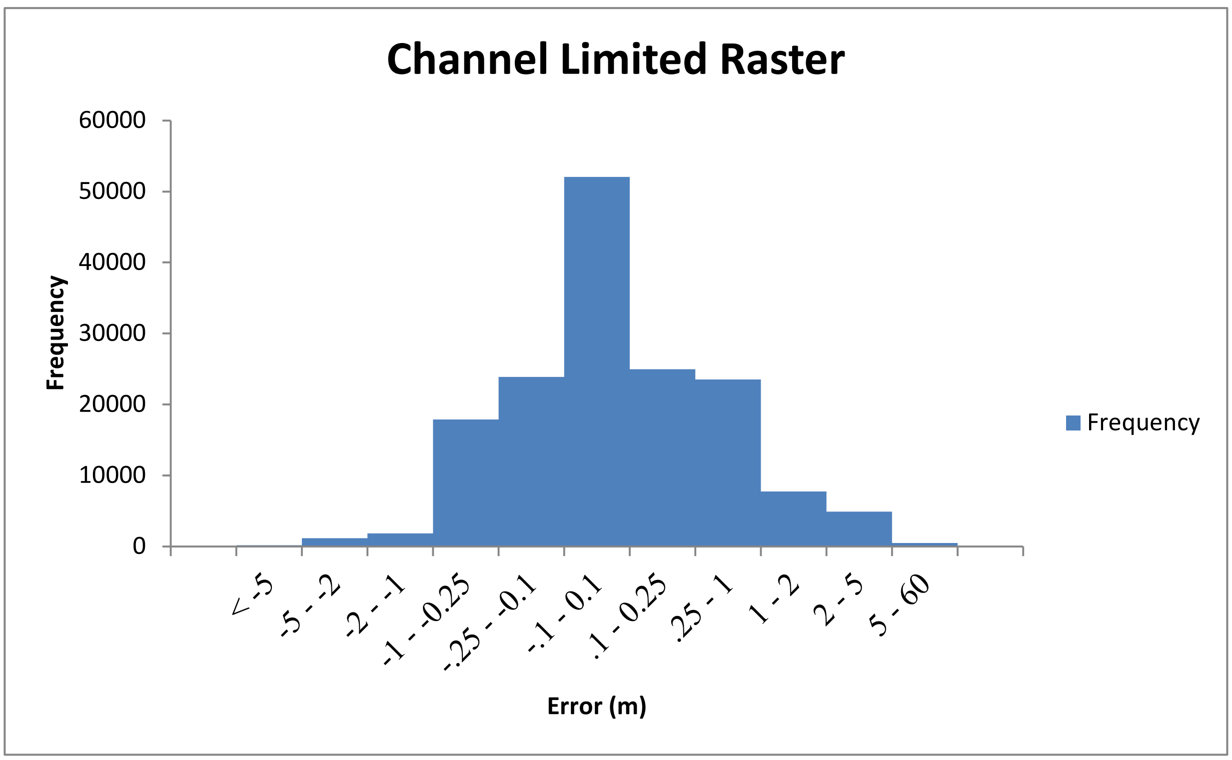


Figure 16: The channel limited raster height error histogram

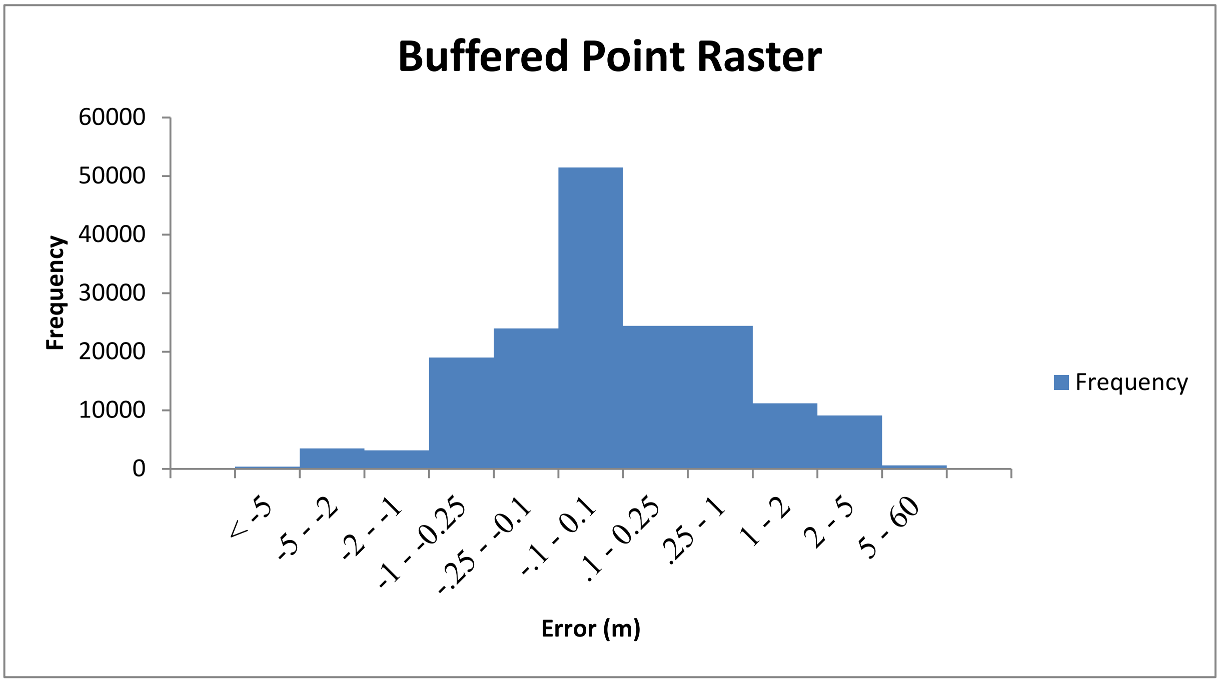


Figure 17: The buffered point raster height error histogram

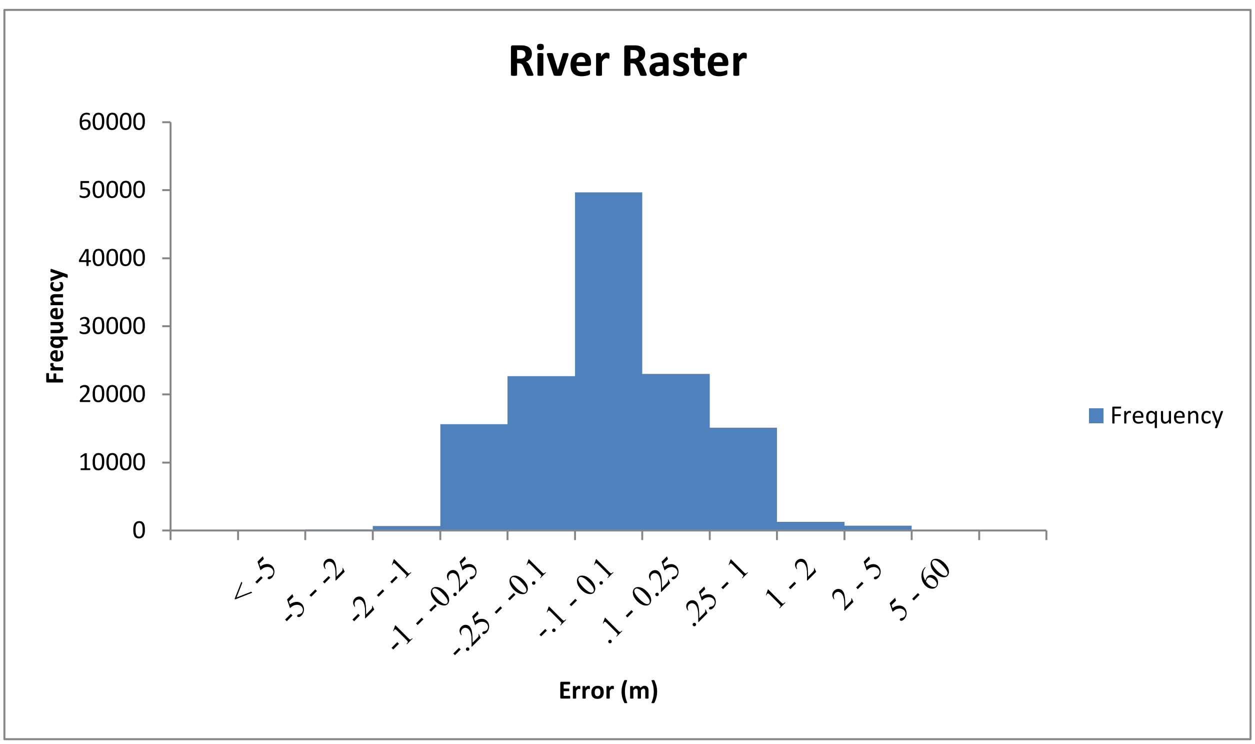


Figure 18: The river raster height error histogram

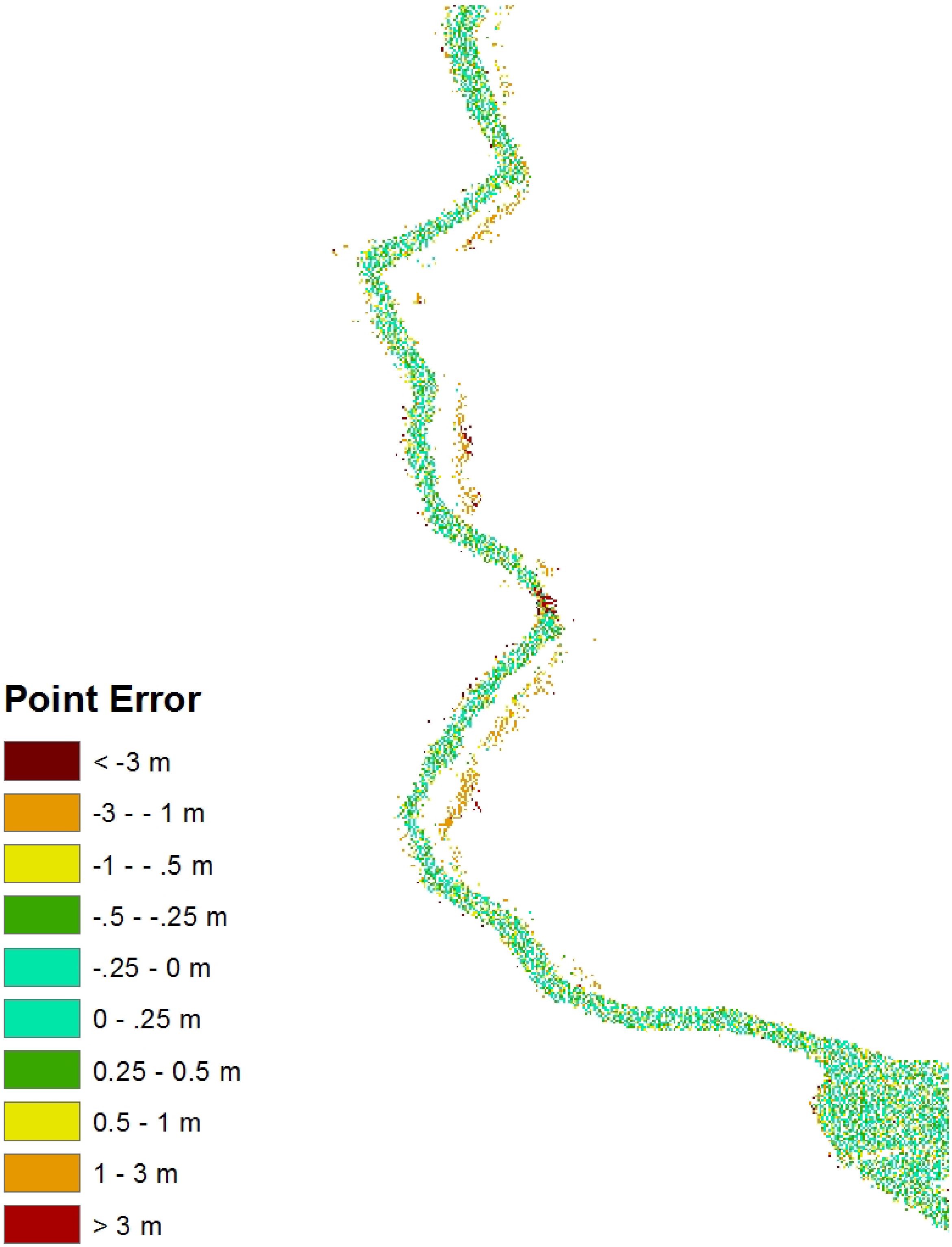


Figure 19: Point raster height error map

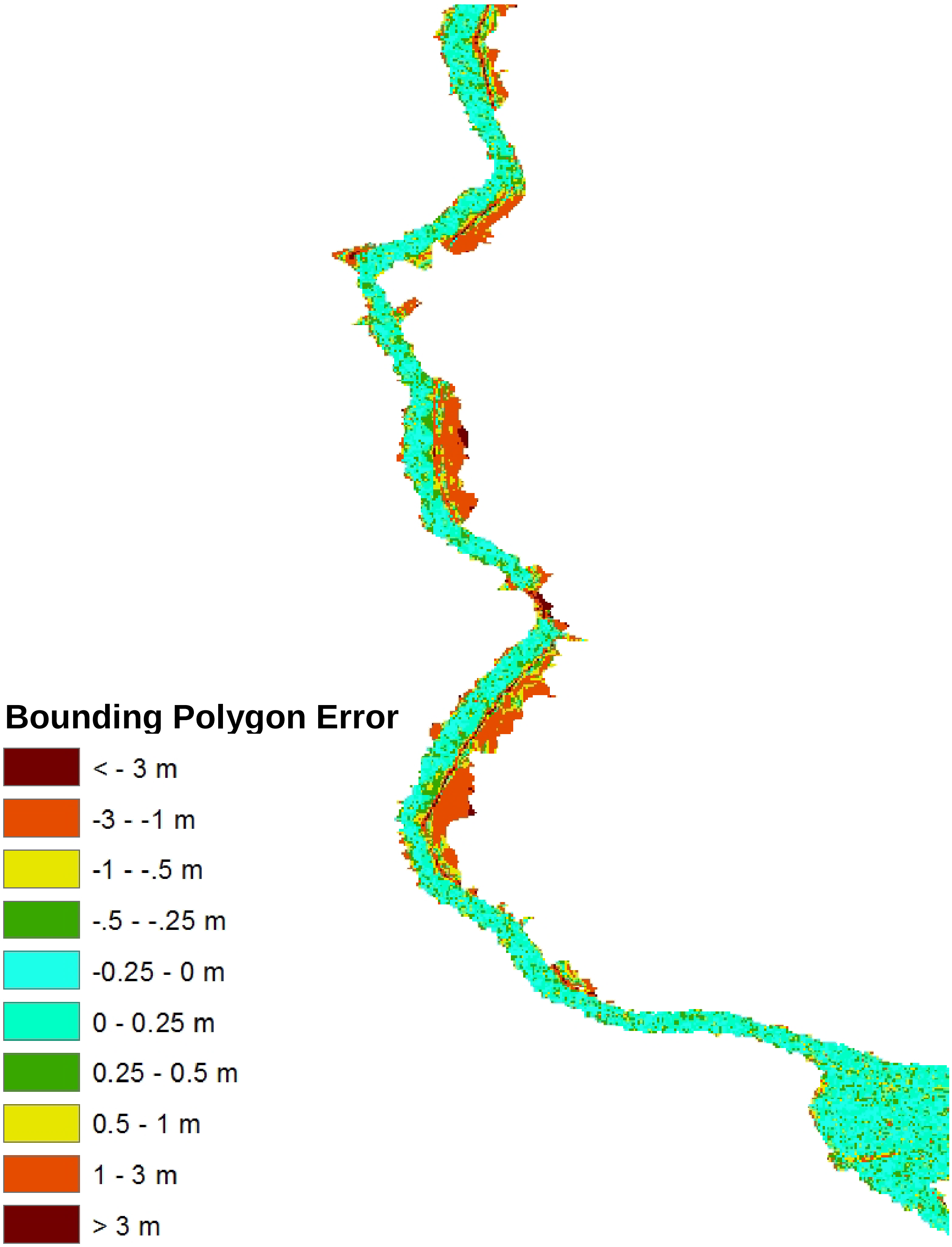


Figure 20: Bounding Polygon height error map

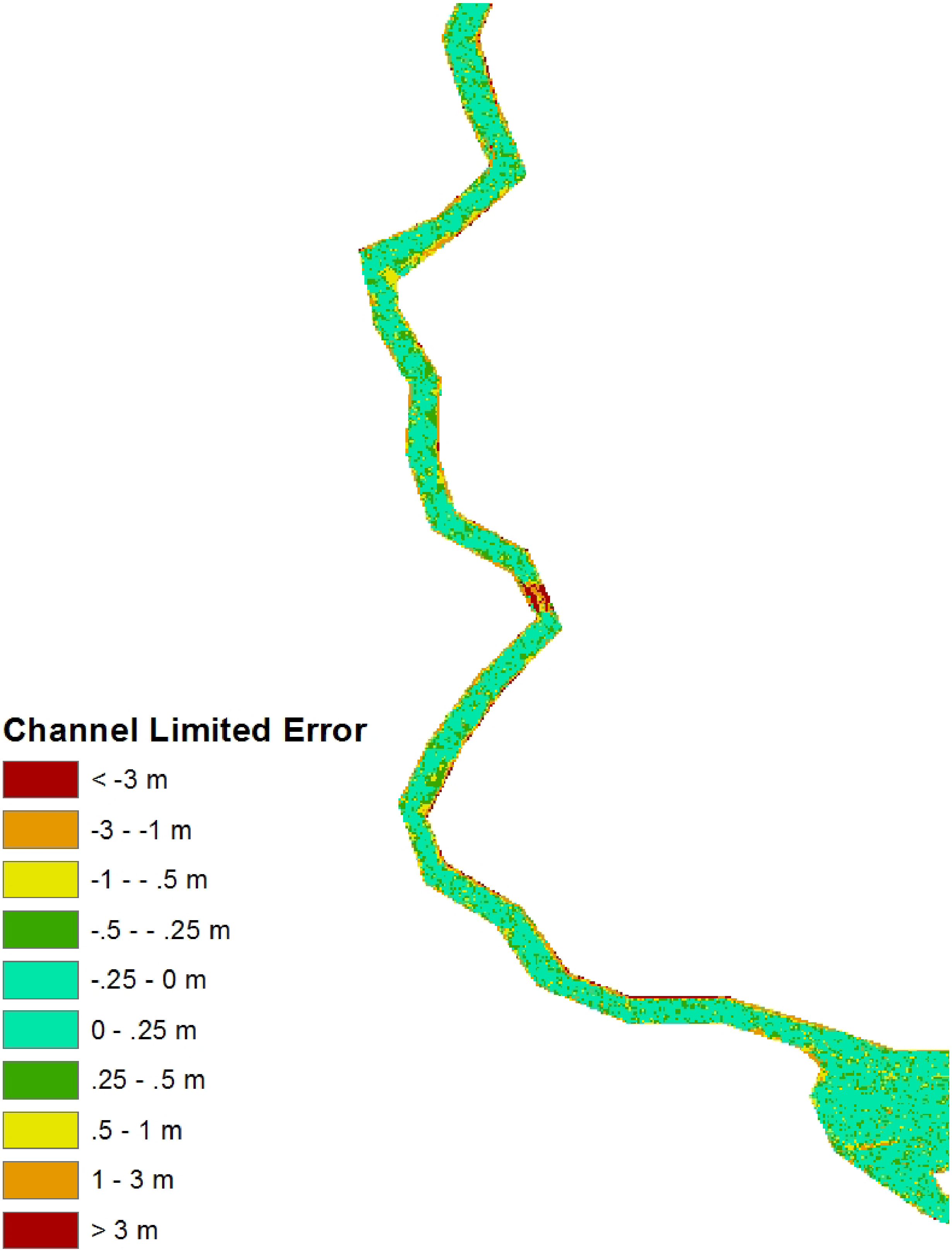


Figure 21: Channel Limited height error map

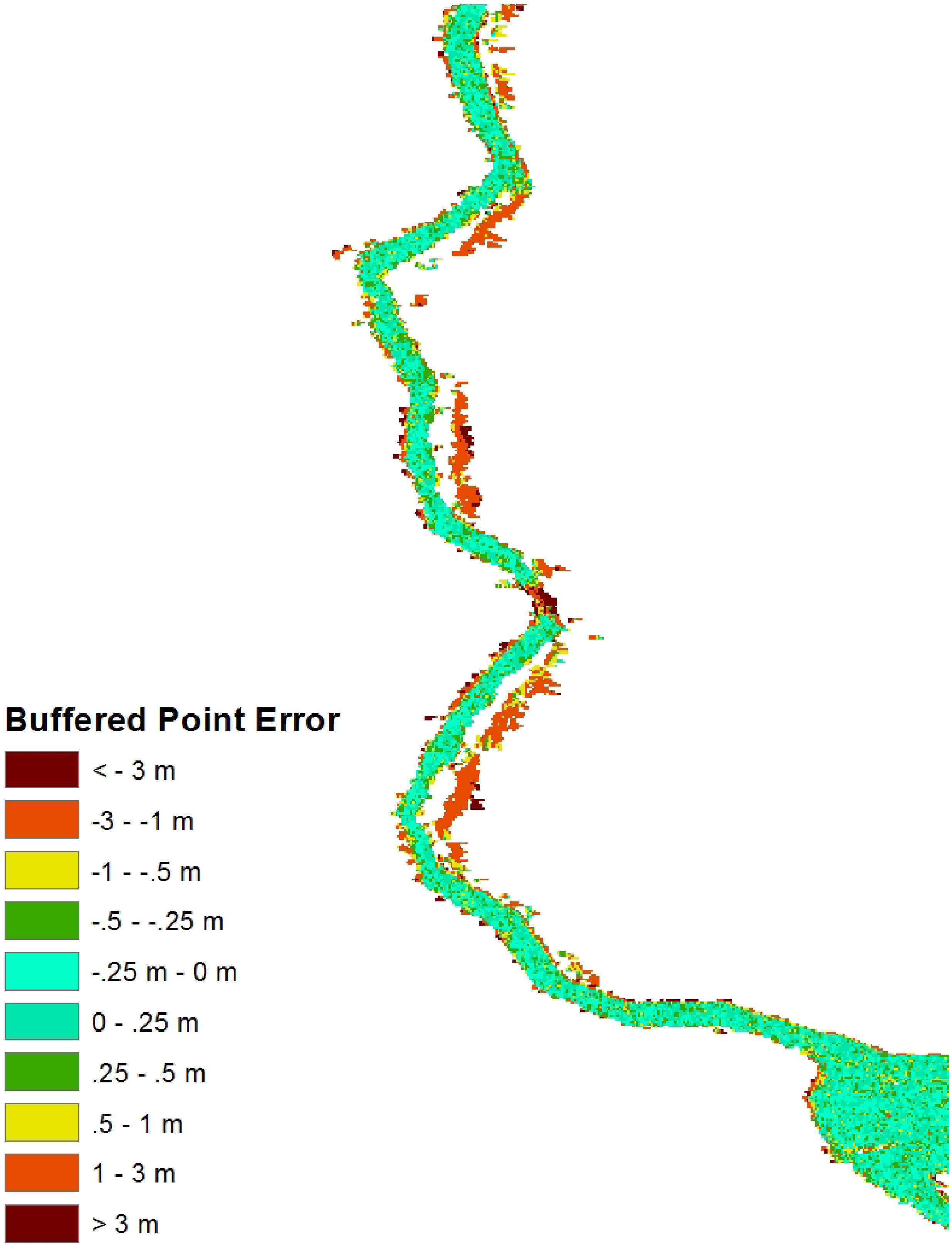


Figure 22: Buffered Point height error map

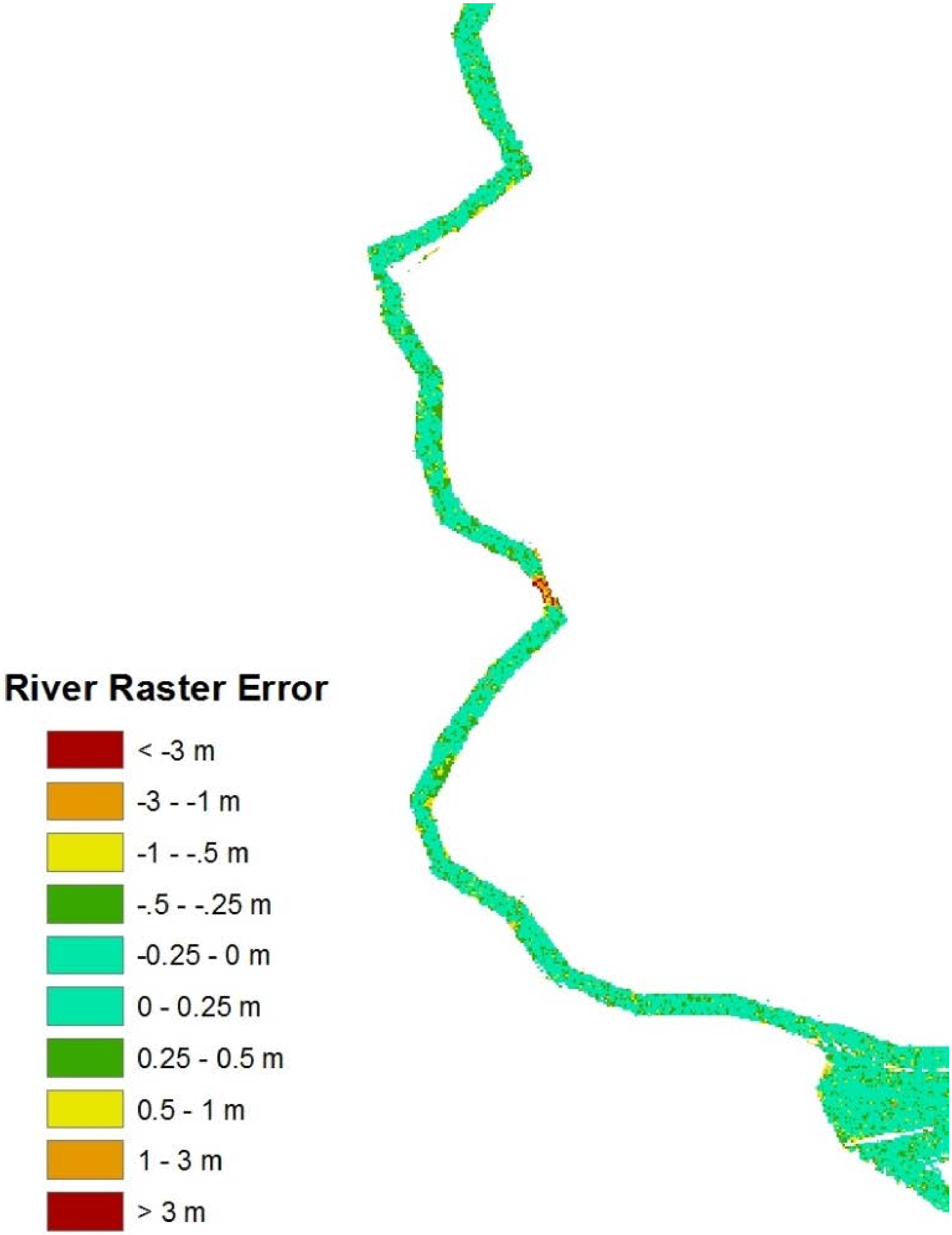


Figure 23: River raster height error map

Width Error

The following figures graphically present the width error. Figures 24 through 26 show the widths of three out of the four raster data sets that were analyzed for height error compared to the reference truth width from the HEC-GeoRAS floodplain output. Figure 27 is used to show how effectively the HEC-GeoRAS tool mapped the inundated area of the HEC-RAS flow simulation, by comparing the RivWidth outputs for the river raster data set with the surveyed cross sectional widths in HEC-RAS.

All variance in calculated width data can be attributed directly to the input files given to RivWidth. When the height maps were created, each raster data set assigned pixels with elevation values based on the type of geometry that was used to define the inundated area from dry land. Although the elevation data is not important, or used for a RivWidth calculation, each technique used to model the river produced different inundated areas to represent the river.

The point raster input to RivWidth had many pixels within the river channel that were missing data. RivWidth unsurprisingly failed to calculate width values for this data set since no continuous body of water could be found. This data set could not be used. The bounding polygon raster agrees with the reference widths for parts of the river that are well defined by SWOT outputs. However this data set over estimated the width of the river in sections where many SWOT outputs fell outside of the river channel. Also because this of how this data set was created, it produced very jagged widths for the whole river.

The channel limited raster input to RivWidth produced width outputs that were very smooth, in comparison to the bounding polygon raster, due to the long straight lines in its parent geometry. However this also produced widths with a consistent bias above the reference width. This is because the floodplain produced by HEC-GeoRAS was based on user specified discharge in

HEC-RAS. This discharge was not a large enough event to completely fill the limits of the channel with water for the entirety of the river. The buffered point raster at first produced very biased data that fell above the reference data; however this is because the process which created the input raster added 100 feet to the river width everywhere there was a point cloud output. Once this was subtracted from the RivWidth output, the buffered point raster produced the best width results when compared to the reference data.

Two noticeable themes were present in all RivWidth outputs. First, each technique produced results that were consistently greater than the reference data. Once again it has been shown that these methods of analysis add area to SWOT outputs. Secondly, all data sets begin to produce unreliable river widths around 55,000 meters downstream. This is because the river channel becomes hard to distinguish from the flooded overbank area in this area, so a better input to RivWidth is required to produce a more accurate result here. After this point the data sets are producing similar widths, they are at varying distances downstream, as if the width profiles are out of phase. This can again be attributed to the inputs provide to RivWidth. Each input had slightly different inundated pixels; this affected how the centerline was calculated for each data set and where the corresponding widths were placed downstream.

The HEC-GeoRAS vs. HEC-RAS width comparison produced encouraging results. The water surface produced by HEC-GeoRAS was already shown to differ from the modeled HEC-RAS water surface by an average of 1.5 meters. The width comparison appears to show similar accuracy. Downstream however the data sets become out of phase again which can be attributed to the centerline calculation in RivWidth. However there is one major discrepancy between the HEC-RAS model and HEC-GeoRAS, it occurs around 57,000 meters downstream, where all data sets produced poor results. Around 65,000 meters, RivWidth fails to calculate a width for the HEC-GeoRAS input, and is programmed to report a width equal to the pixel size of the input file. RivWidth assumes that the river becomes very thin here and is not properly depicted in the input file. This requires more examination because the river maintains a width here of over 100 meters.

It appears as though the experimental techniques report width values that are closer to the HEC-RAS cross section widths, each technique over estimated river width, while HEC-GeoRAS consistently underestimates the widths. This phenomenon is most likely just chance, since the rasters were created off of the HEC-GeoRAS floodplain and never interacted with HEC-RAS data.

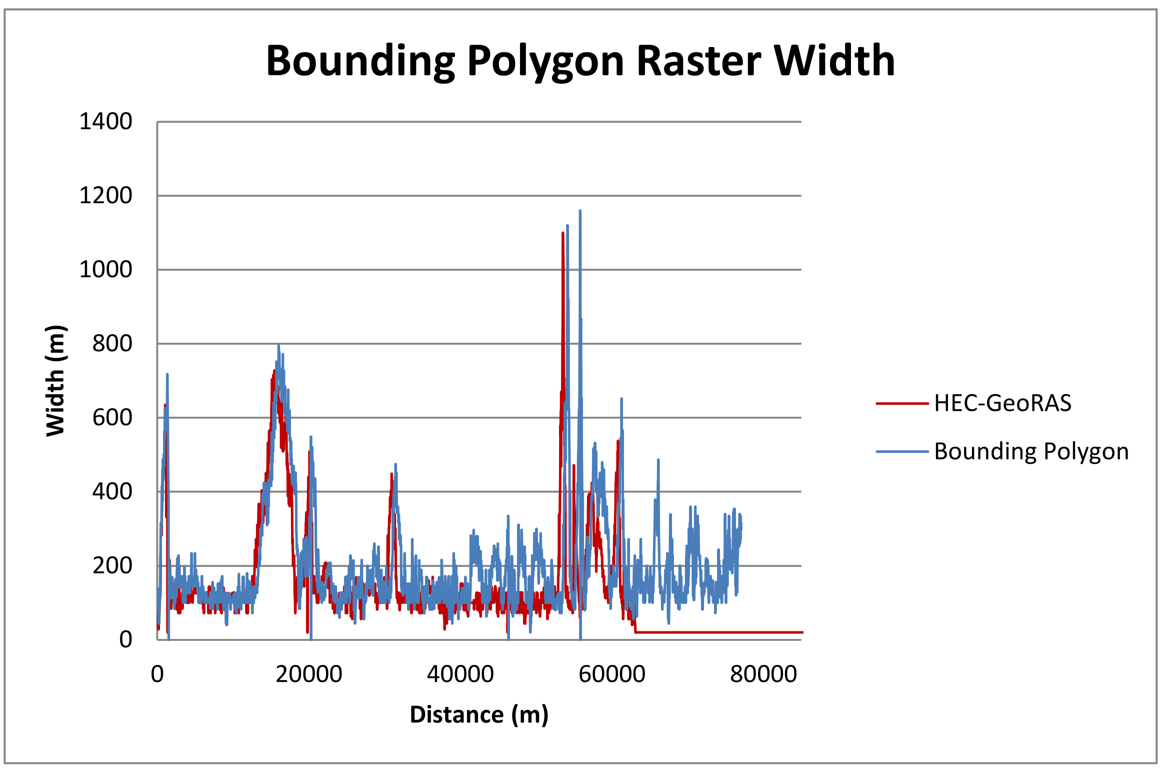


Figure 24: Comparison of the Bounding polygon width outputs to reference data

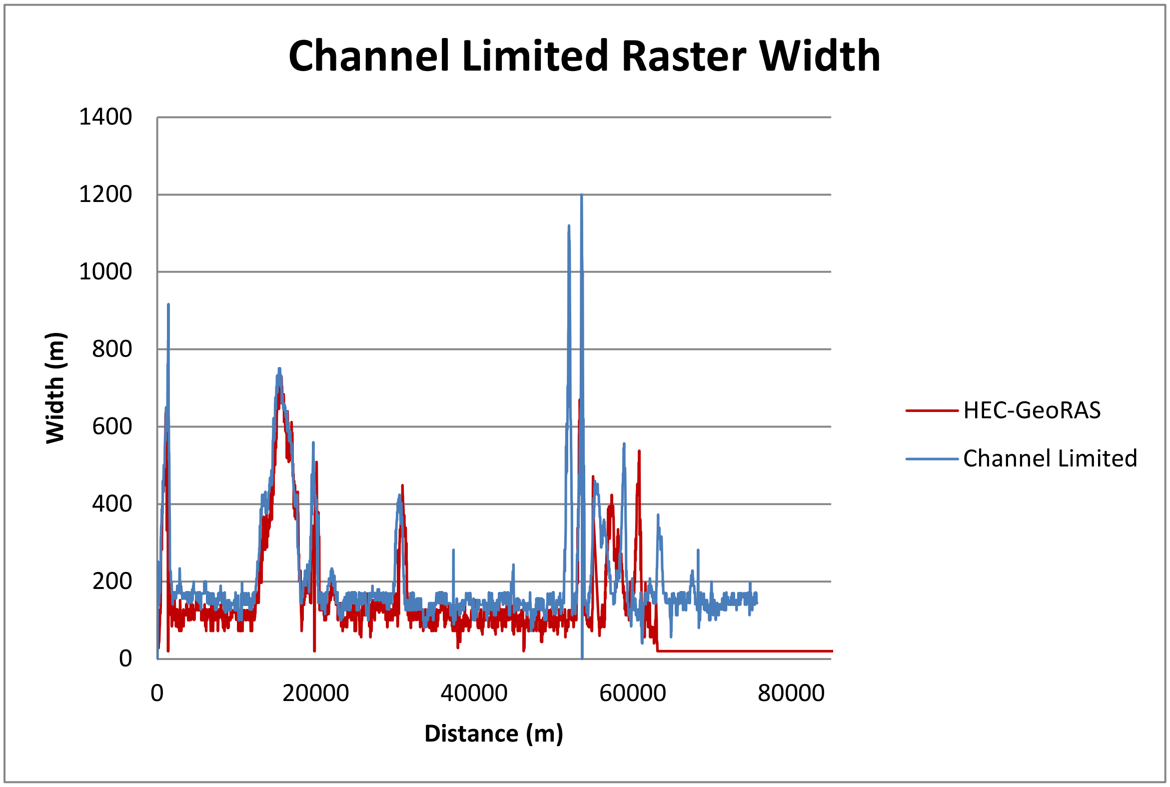


Figure 25: Comparison of the channel limited raster width outputs to reference data.

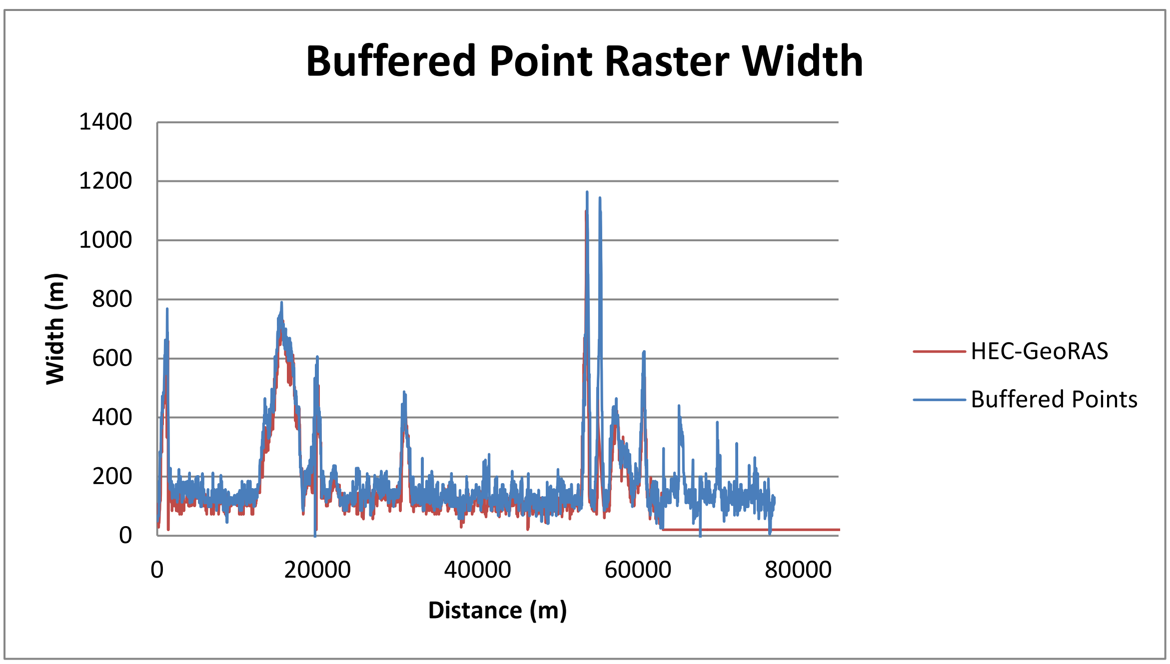


Figure 26: Comparison of the buffered point raster width outputs to reference data.

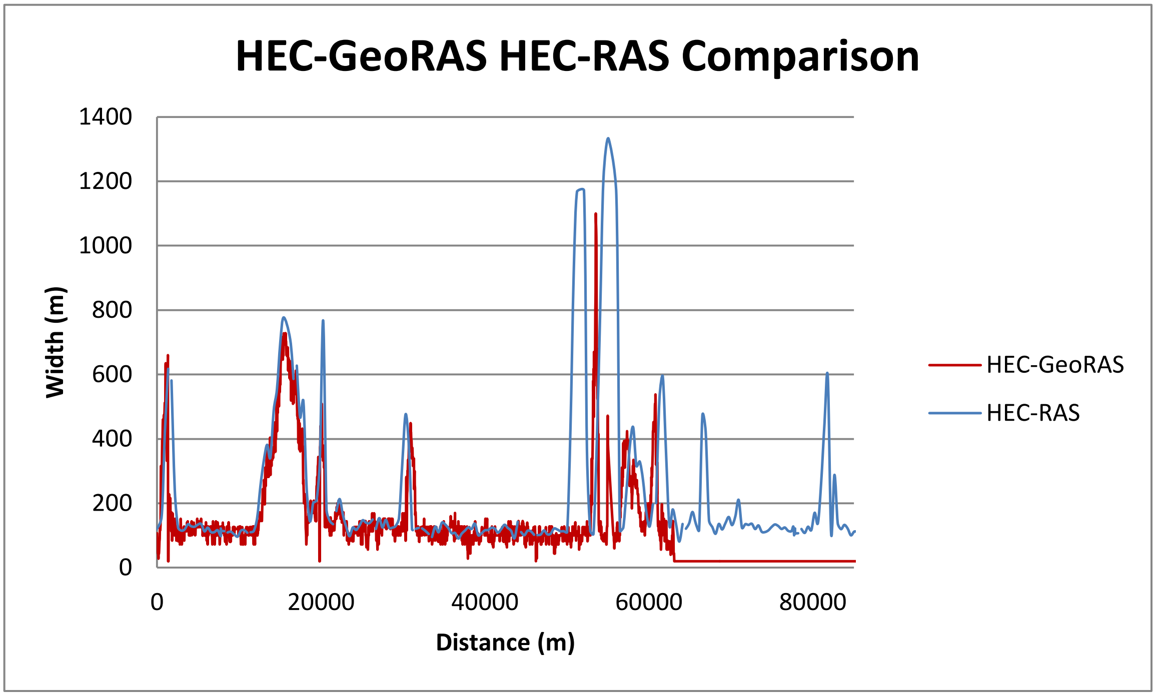


Figure 27: Comparison of HEC-GeoRAS to HEC-RAS modeled data.

Slope Error

Slope data was calculated for each of the height raster data sets. Elevation data for points found along the river's centerline were plotted against distance and compared to the reference water surface slope. This reference water surface slope was created in the same manner as the experimental surface slopes, only using the HEC-GeoRAS floodplain as its source for elevation data. The slope information inside the HEC-RAS model was also obtained to once again verify HEC-GeoRAS results. This data is represented in Figures 28 through 31, and combined in Table 4.

These outputs once again verify HEC-GeoRAS is working correctly showing less than 3 millimeter per kilometer difference in slope from the HEC-RAS model. The worst slope obtained from an input raster was the point raster, but even this error was only about 5 millimeters per kilometer from the reference surface. All other methods are almost identical in slope values.

All methods produced nearly the same slope because the river centerline was used to calculate the slope of each raster surface. The centerline elevation was an average of a channel bound zone, averaging only pixels that were known to be water. As discussed above, the height errors greatest in magnitude occurred near the channel limits and outside of the channel. The interpolated surfaces representing the river channel in each of these different height maps is very similar. Since each of the methods produced heights measurements for the entire study area, the same centerline distance was also used in the slope calculations.

The slope profile for the study area shows an almost constant decline in water surface which is to be expected, however around 11,000 meters the river does flatten out, corresponding to a decrease in water depth (Figure 32), before continuing to decline. Around 65000 meters downstream, there is confusing height data that distorts the slope calculation.

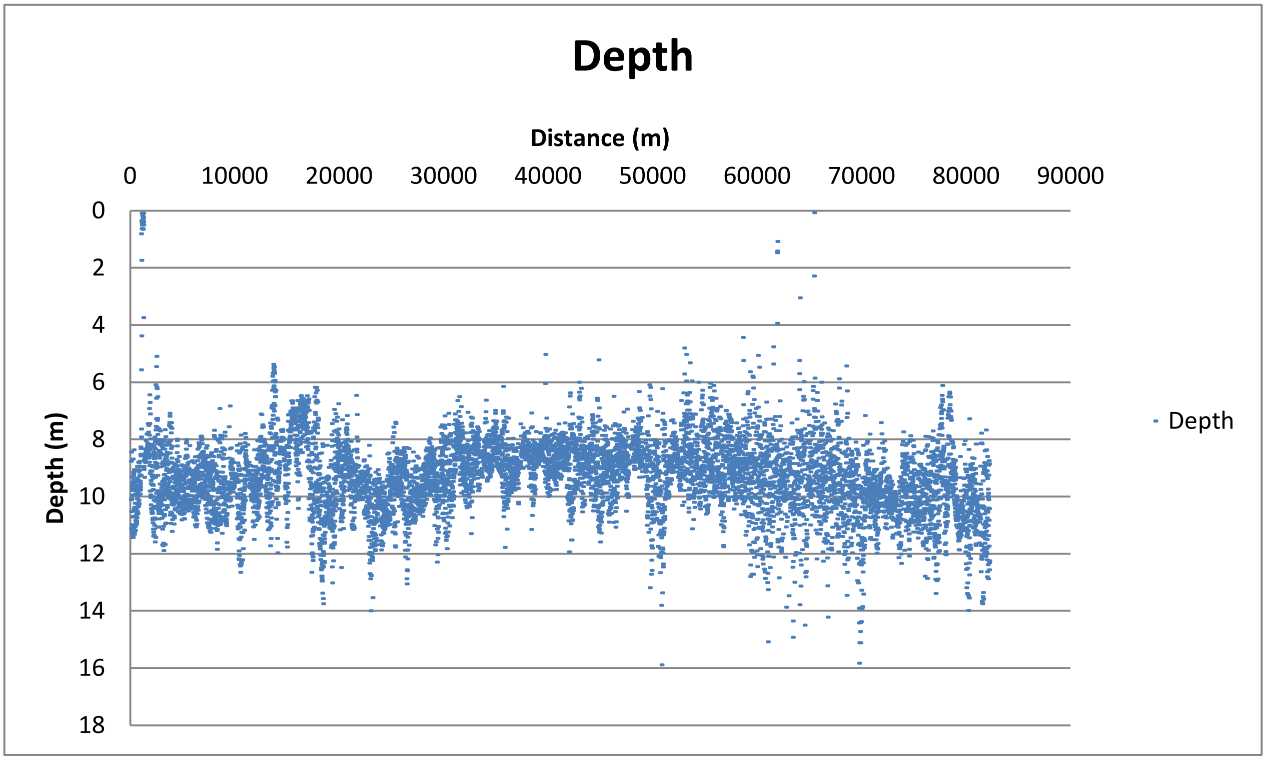


Figure 28: Depth profile for the study area

Method	Slope
Point Raster	-9.951 cm/km
Bounding Polygon Raster	-9.595 cm/km
Channel Limited Raster	-9.592 cm/km
Buffered Points Raster	-9.603 cm/km
HEC-GeoRAS	-9.444 cm/km
HEC-RAS	-9.252 cm/km

Table 4: This table lists slope results next to their corresponding source.

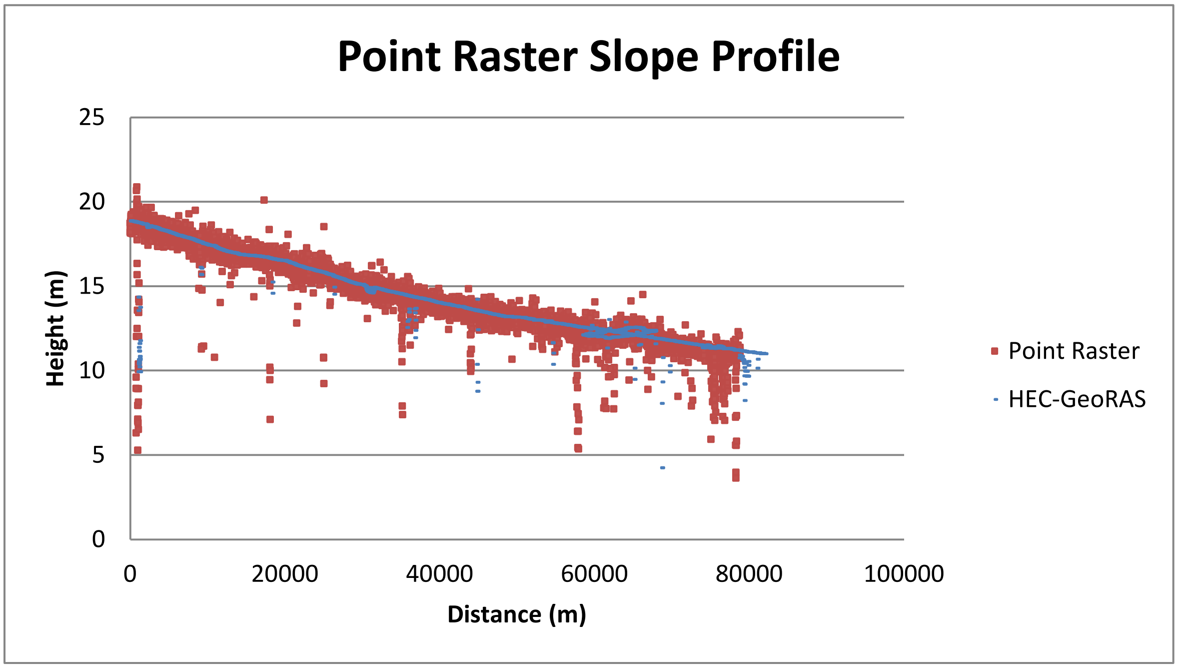


Figure 29: Slope profile for Point Raster

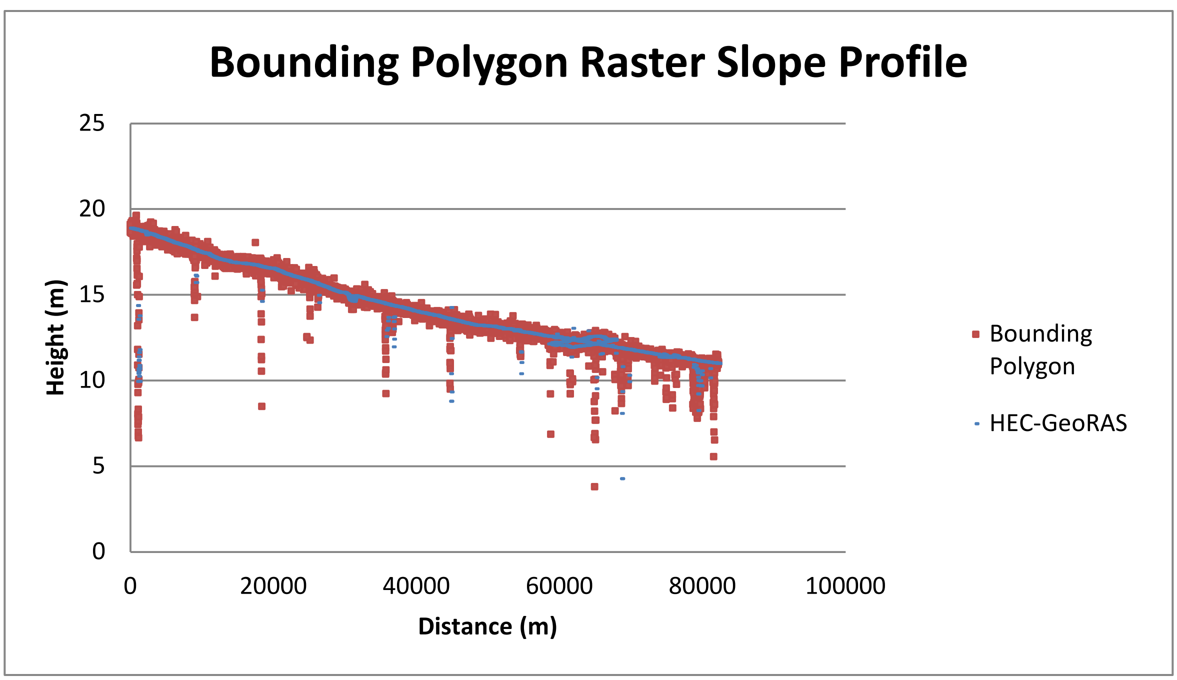


Figure 30: Slope profile for Bounding Polygon Raster

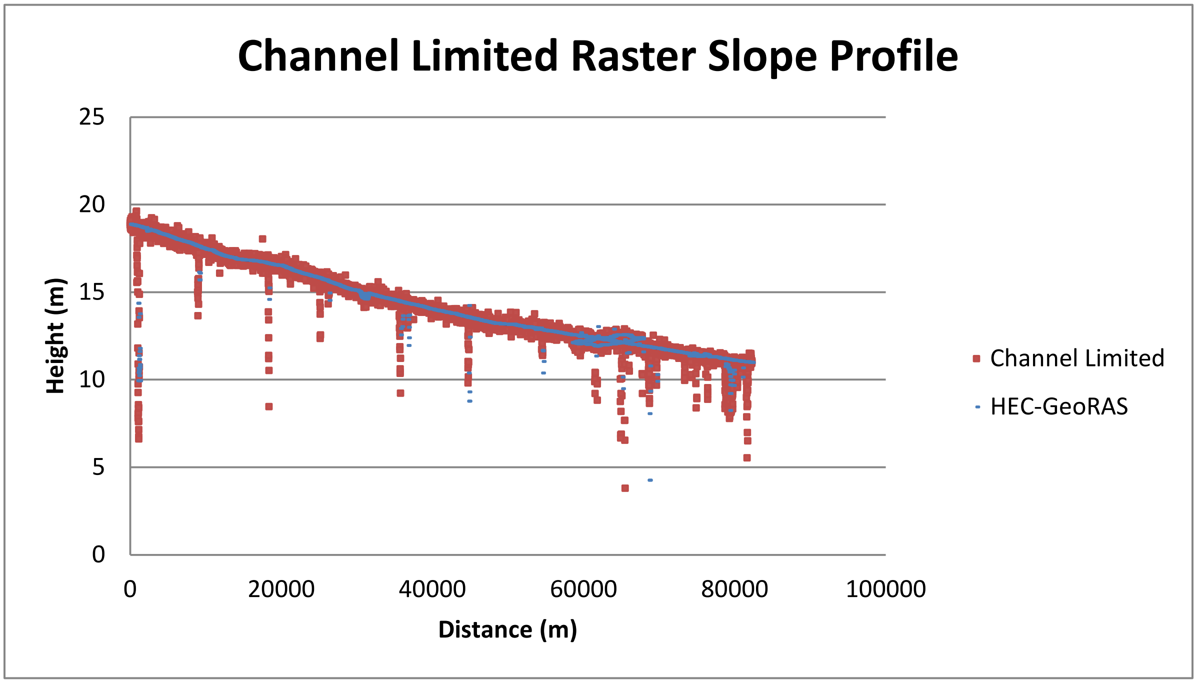


Figure 31: Slope profile for Channel Limited Raster

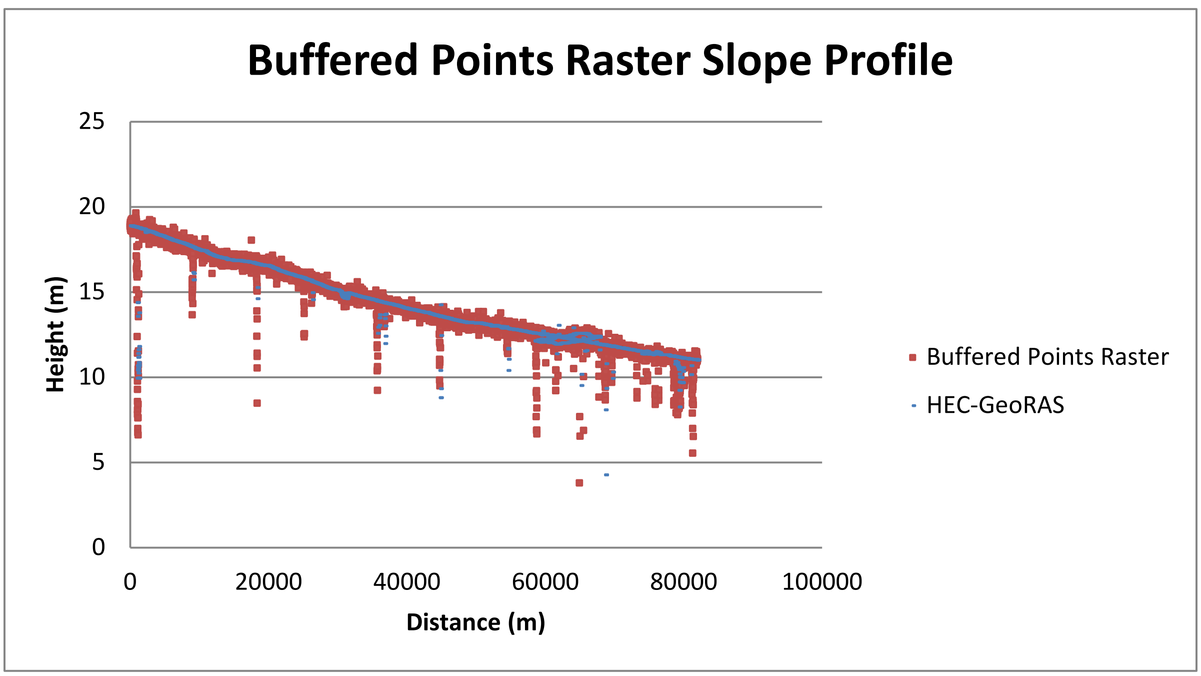


Figure 32: Slope profile for Buffered Point Raster

Discharge Error

The average cross sectional flow area and roughness coefficient used to calculate the experimental discharges for the five sub reaches, are reach averaged values taken from HEC-RAS cross section data. A reach averaged discharge for the HEC-RAS model was calculated and used as the accepted discharge value when calculating mean percent error and absolute mean percent error for each of the experimental discharges. This data is shown in Table 5. Each of the raster data sets that produced width and slope estimates were then used to calculate discharge. The calculated discharges were then compared to the reach averaged HEC-RAS model discharges and the errors for each raster data set were calculated.

It became very clear that these errors could be reduced by eliminating poor SWOT height measurements. In order to filter the data, a linear trend line was found of the original observations for each of the five sub reaches. The downstream distance of each SWOT measurement was used as the X coordinate to calculate the corresponding Y value (Height). The difference between the SWOT observations and the linear estimates, known as residuals, were found and the standard deviation of the residuals was calculated. Any SWOT measurement that had a residual greater than twice the standard deviation of all residuals was not used for slope calculation. Discharge was then recalculated with the new slope estimations (Table 6) and the mean percent error and absolute mean percent error for each data set, other than the river raster, drastically reduced when compared to the unfiltered discharge estimation errors. The filtered discharge estimation errors are listed in Table 7.

Each technique was shown to both underestimate and overestimate discharge depending on the sub reach, and this lowered the mean percent error. The buffered point raster data set produced the most accurate results with a mean percent error of -.5%; however the other experimental techniques were close behind with mean percent errors of around -4%, while the river raster data set produced the worst discharge estimates with a mean percent error of over 13%. When calculating mean percent error, with absolute differences instead of positive or negative differences, the river raster maintained the worst error, almost 16.5%, the other techniques produced errors close to 7%.

To further evaluate the discharge error, the true HEC-RAS slope values were used for all experimental methods and errors due to width calculation were found. For each technique the mean percent errors and the absolute mean percent errors experienced negligible changes (Table 8). Discharges were also calculated using the experimental slope values for each raster data set, and true HEC-RAS widths for each sub reach to determine the mean percent error due to errors in width estimation. This data is collected in (Table 9). Comparing Table 8 and Table 9, the data shows that the height estimations and resulting slope estimations for the five sub reaches derived from the simulator output, are better than the width estimations based off of the techniques used to emulate the river's true geometry.

HEC - RAS Model Data	Discharge (m^3/s)	n	Cross Sectional Area (m^2)	Width (m)	Slope
Reach 1	976.931	0.035	935.545	168.456	0.00013582
Reach 2	1099.413	0.035	1567.656	385.149	0.00009271
Reach 3	1031.106	0.035	943.382	162.625	0.00014040
Reach 4	774.625	0.035	771.146	126.005	0.00011042
Reach 5	665.304	0.035	820.5231	156.763	0.00008862

Table 5: Calculated reach averaged discharge for HEC-RAS data

Reach 1	Discharge (m^3/s)	n	Cross Sectional Area (m^2)	Width (m)	Slope
Bounding Polygon Raster Product	997.7605926	0.035	935.545	168.444	0.00014166
Channel Limited Raster Product	868.5456827	0.035	935.545	202.836	0.00013752
Buffered Polygon Raster Product	952.8660467	0.035	935.545	175.165	0.00013612
River Raster Product	1126.201446	0.035	935.545	144.971	0.00014775
Reach 2					
Bounding Polygon Raster Product	1122.349442	0.035	1567.656	347.142	0.00008412
Channel Limited Raster Product	1057.79811	0.035	1567.656	376.18	0.00008317
Buffered Polygon Raster Product	1060.292208	0.035	1567.656	374.049	0.00008293
River Raster Product	1141.250386	0.035	1567.656	331.709	0.00008186
Reach 3					
Bounding Polygon Raster Product	1065.28275	0.035	943.382	153.976	0.00013933
Channel Limited Raster Product	1036.047243	0.035	943.382	164.661	0.00014412
Buffered Polygon Raster Product	1091.329111	0.035	943.382	152.432	0.00014427
River Raster Product	1224.979606	0.035	943.382	127.235	0.00014286
Reach 4					
Bounding Polygon Raster Product	637.1872168	0.035	771.146	155.561	0.00009895
Channel Limited Raster Product	665.5012454	0.035	771.146	146.391	0.00009954
Buffered Polygon Raster Product	648.2641111	0.035	771.146	152.012	0.00009932
River Raster Product	708.9000141	0.035	771.146	128.544	0.00009497
Reach 5					
Bounding Polygon Raster Product	612.704736	0.035	820.5231	177.782	0.00008889
Channel Limited Raster Product	715.7011999	0.035	820.5231	141.284	0.00008928
Buffered Polygon Raster Product	759.68273	0.035	820.5231	128.922	0.00008903
River Raster Product	904.9997278	0.035	820.5231	99.136	0.00008901

Table 6: Table displaying all reach-averaged discharges for different filtered raster techniques

Discharge Error	Reach 1 (m^3/s)	Reach 2 (m^3/s)	Reach 3 (m^3/s)	Reach 4 (m^3/s)	Reach 5 (m^3/s)	Mean Percent Error	Absolute Mean Percent Error
Bounding Polygon Raster	20.829	22.937	34.177	-137.438	-52.599	-3.623 %	6.636 %
Channel Limited Raster	-108.385	-41.614	4.941	-109.124	50.397	-4.183 %	7.404 %
Buffered Point Raster	-24.065	-39.120	60.223	-126.361	94.379	-0.462 %	8.472 %
River Raster	149.270	41.838	193.874	-65.725	239.696	13.086 %	16.480 %

Table 7: Errors in discharge of different filtered data sets

Discharge Error (True Slopes)	Reach 1 (m^3/s)	Reach 2 (m^3/s)	Reach 3 (m^3/s)	Reach 4 (m^3/s)	Reach 5 (m^3/s)	Mean Percent Error	Absolute Mean Percent Error
Bounding Polygon Raster	0.046	78.849	38.259	-101.520	-53.531	-2.053 %	6.408 %
Channel Limited Raster	-113.771	17.406	-8.517	-73.696	47.747	-2.645 %	6.149 %
Buffered Polygon Raster	-25.107	21.644	45.468	-91.084	92.630	1.195 %	6.926 %
River Raster	102.846	115.118	183.281	-10.234	237.711	14.839 %	15.165 %

Table 8: Errors in discharge of filtered data sets calculated with true HEC-RAS slopes

Discharge Error (True Widths)	Reach 1 (m^3/s)	Reach 2 (m^3/s)	Reach 3 (m^3/s)	Reach 4 (m^3/s)	Reach 5 (m^3/s)	Mean Percent Error	Absolute Percent Error
Bounding Polygon Raster	20.782	-52.171	-3.937	-41.335	1.013	-1.637 %	2.549 %
Channel Limited Raster	6.095	-58.101	13.571	-39.152	2.473	-1.605 %	2.530 %
Buffered Polygon Raster	1.069	-59.591	14.132	-39.978	1.535	-1.774 %	2.458 %
River Raster	42.002	-66.334	8.994	-56.234	1.462	-1.580 %	3.737 %

Table 9: Errors in discharge of filtered data sets calculated with true HEC-RAS widths

Conclusion

This study shows that simulated SWOT results can produce height estimated of within 2 to 22 centimeters of accuracy, and slope estimates of within 2 to 7 centimeters of accuracy for an 80 kilometer study area. Discharges estimations for 5 sub reaches fell within -12% to 4% when using unfiltered simulator outputs, and -4% to 13% when filtered. Measurement accuracy is heavily dependent on what techniques are used to analyze the SWOT simulation results.

The discharge estimations are misleading at first. The river raster data set was thought to represent true river geometry. It was based on the floodplain generated by HEC-GeoRAS. Therefore it was assumed that this data set would produce discharge estimations with the lowest errors. However it was shown that HEC-GeoRAS produces a water surface that is 1.5m on average lower than the HEC-RAS model resulting in river widths that are consistently less than the HEC-RAS model. The misrepresentation and constant bias of HEC-RAS data with HEC-GeoRAS leads to miscalculations in river discharge. Therefore, the river raster data set should not be used when determining a reasonable estimation of the accuracy of SWOT measurements or the ability to use these measurements to calculate discharge.

The other methods produce discharge estimations within -4% to -.5% of the true discharge. These methods are more accurate mainly because of the quality of their river widths. Each technique was shown to produce widths greater than HEC-GeoRAS, unintentionally closer to the true width values found in HEC-RAS. Overall the buffered point raster data set produced a mean percent error of -.5%, and an absolute mean percent error of 8.5%. In future work with simulated SWOT results, the buffered point technique should use used, if the SWOT measurements are intended to be treated as a point cloud, until a more robust method can be formulated.

The current requirement for SWOT height measurement is to have 10 centimeter per kilometer error. This number is achieved by the point raster and the river raster, the other raster data sets, with filtering, can become more accurate to meet this requirement. The current requirement for slope measurement is 1 cm per kilometer, over a 10 kilometer reach. Every technique produced slope errors less than 1 cm per kilometer for each sub reach, excluding sub reach 1. In this sub reach there is significant height estimation error and this reflects in the slope estimation accuracy. After filtering all slope estimations for all sub reaches met this requirement. The requirement for discharge is less than 20% error. When considering either mean percent error or absolute mean percent error, all discharge estimations are within this requirement.

Suggestions for Future Work

Three main suggestions for future work are, first to repeat the buffered point method for multiple flood events over the same study area. This work was done on one flow simulation. Although it produced useful information on what can be obtained from SWOT outputs, SWOT is being designed to be a multiple year mission. Evaluating SWOT's performance over a series of days or months will prove more useful and be more appropriate than a single event study.

The second suggestion for future work is to produce SWOT output for the other orbit files that overlap this study area DEM. More SWOT outputs will allow the interpolated water surface to be more accurate locally around each point cloud point. Also other orbit passes may provide further insight into why some SWOT measurements fall outside of the river channel. If this is due to layover, as expected, then perhaps other orbit passes will provide insight on how to mitigate these observations.

Lastly, understanding the transition from HEC-RAS to HEC-GeoRAS needs to be further explored. The HEC-GeoRAS water surface was significantly lower than the HEC-RAS model and although this misrepresentation did not impede this study, it would be beneficial for this to be resolved in future work.

Works Cited

ESRI (Environmental Systems Resource Institute). ArcGIS Desktop Help 10.0., http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html?TopicName=welcome., 2010

Pearl, E. Rodriguez, E. Moeller, D. McAdams, M. Johnson, M. Konstantinos, A. Arumuganm, D. ., SWOT Hydrology Simulator Users Guide., version 2.5., FEBRUARY 2013

Roads, J. Lawford, R. Bainto, E. Berbery, E. Chen, S. Fekete, B. Gallo, K. Grundstein, A. Higgins, Kanamitsu, W. Krajewski, M. Lakshmi, W. Leathers, V. Lettenmaier, D. Luo, L. Maurer, E. Meyers, T. Miller, D. Mitchell, K. Mote, T. Pinker, R. Reichler, T. Robinson, D. Robock, A. Smith, J. Srinivasan, G. Verdin, K. Vinnikov, K. Vonder Haar, T. Vörösmarty, C. Williams, S. Yarosh, E., GCIP Water and Energy Budget Synthesis (WEBS), J.Geophys.Res., 2003

Shiklomanov, A.I., R.B.Lammers, and C.J.Vörösmarty, Widespread decline in hydrological monitoring threatens Pan-Arctic research, Eos Trans., AGU, 83, 13–16, 2002.

Tamlin M. Pavelsky and Laurence C. Smith, RivWidth: A Software Tool for the Calculation of River Widths From Remotely Sensed Imagery., IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 5, NO. 1, JANUARY 2008

US Army Corps of Engineers Institute for Water Resources Hydrologic Engineering Center, HEC-RAS River Analysis System User’s manual, version 4.1., JANUARY 2010

US Army Corps of Engineers Institute for Water Resources Hydrologic Engineering Center, HEC-GeoRAS GSI Tools for Support of HEC-RAS using ARCGIS 10 User’s Manual, version 10., MAY 2012